RNA EDITING IN TRYPANOSOMES: SUBSTRATE RECOGNITION AND ITS INTEGRATION TO RNA METABOLISM

A Dissertation

by

ALFREDO J. HERNANDEZ

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

December 2010

Major Subject: Biochemistry

RNA Editing in Trypanosomes: Substrate Recognition and its Integration to RNA Metabolism

RNA EDITING IN TRYPANOSOMES: SUBSTRATE RECOGNITION AND ITS INTEGRATION TO RNA METABOLISM

A Dissertation

by

ALFREDO J. HERNANDEZ

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Jorge Cruz-Reyes
Committee Members,	Ryland Young
	David Peterson
	Xiuren Zhang
Head of Department,	Gregory Reinhart

December 2010

Major Subject: Biochemistry

ABSTRACT

RNA Editing in Trypanosomes: Substrate Recognition and its Integration to RNA Metabolism. (December 2010)

Alfredo J. Hernandez, B.S., The University of Texas- Pan American

Chair of Advisory Committee: Dr. Jorge Cruz-Reyes

RNA editing in trypanosomes is the post-transcriptional insertion or deletion of uridylates at specific sites in mitochondrial mRNAs. This process is catalyzed by a multienzyme, multisubunit complex through a series of enzymatic cycles directed by small, trans-acting RNA molecules. Despite impressive progress in our understanding of the mechanism of RNA editing and the composition of the editing complex, fundamental questions regarding RNP assembly and the regulation of catalysis remain.

This dissertation presents studies of RNA-protein interactions between RNA editing complexes and substrate RNAs and the determination of substrate secondary structural determinants that govern them. Our results suggest that substrate association, cleavage and full-round editing by RNA editing complexes in vitro obey hierarchical determinants that increase in complexity as editing progresses and we propose a model for substrate recognition by RNA editing complexes.

In addition, this dissertation also presents the characterization of a novel mitochondrial RNA helicase, named REH2 and its macromolecular interactions. Our

data suggest that REH2 is intimately involved in interactions with macromolecular complexes that integrate diverse processes mediating mitochondrial gene expression.

These results have implications for the mechanism of substrate RNA recognition by RNA editing complexes as well as for the integration of RNA editing to other facets of mitochondrial RNA metabolism.

DEDICATION

I dedicate this dissertation to my family, friends and to the memory of those no longer here.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Jorge Cruz-Reyes, for giving me a chance way back when and for all the things I have learned in his lab. Thanks to Dr. Bhaskar Reddy Madina, the newbies, Damien Terry and Vikas Kumar, in addition to former members of the lab for their friendship and camaraderie.

I would also like to thank the members of my doctoral committee, Dr. Young, Dr. Peterson and Dr. Zhang, for their guidance and support. In addition, I would like to thank Dr. Shippen for seeing me as her "unofficial" graduate student.

Many thanks go to the staff of the Department of Biochemistry and Biophysics: to Pat Swigert, who always tried, though not always could, to keep me in line and get me to submit stuff by the deadline. Thanks to Daisy Wilbert, Tillie Rausch and Juanita Withem for their cheerful, earnest help and to Sherri Coronado for helping me with tuition stuff when I was out of the country.

I also want to extend my gratitude to the National Science Foundation and everyone at TAMU BTD, which provided funding for my first two years of graduate school.

Thanks to my family for their love: To my mother and father for their encouragement and support; to my brother and sister for foolishly looking up to me; to my grandfather, for being an exemplar of what I should aspire to be; and last, but not least, thanks to my wife, for everything you do and everything you are. You are everything to me.

TABLE OF CONTENTS

CHAPTER Page

CHAPTER Page

LIST OF FIGURES

FIGURE Page

FIGURE Page

LIST OF TABLES

CHAPTER I

INTRODUCTION

Kinetoplastid organisms are considered to be among the earliest-branching eukaryotes. They are a group of unicellular, flagellate protozoa that share a characteristic subcellular structure, the kinetoplast (1,2). The genus Trypanosoma is of special interest because among its members are the "Tritryps": *Trypanosoma brucei*, *Trypanosoma cruzi*, and *Leishmania major*, obligate parasites agents of devastating human and livestock diseases that are transmitted by insect vectors (Table 1) (3). *Trypanosoma brucei* is the causative agent of African sleeping sickness in humans and Nagana (or wasting disease) in cattle. The disease was formally described in 1734 by the English naval surgeon John Atkins, but accounts of the disease exist as early as 1200 A.D (4). The disease is endemic to sub-saharan Africa and infects approximately 80,000 people annually (5). If left untreated, the outcome is invariably death, although the existing chemotherapeutic treatments are themselves toxic and unsafe. *Trypanosoma cruzi* causes Chaga's disease, a chronic infection that can result in severe cardiomyopathy (6). This disease is endemic in central and south America and affects approximately 8–9 million infected individuals, causing ca. 14,000 deaths annually (5).

This dissertation follows the style of *The Journal of Biological Chemistry*.

Leishmania parasites are found throughout the world, infecting an estimated 12 million individuals (5). Infection with these parasites exhibit a variety of debilitating and disfiguring cutaneous disease as well as other symptoms that can be fatal (5).

Table 1. Diseases Caused by Trypanosomes.

Disease	Human African Trypanosomiasis	Chagas Disease	Leishmaniases
Causative organism	*Trypanosoma brucei gambiense* *Trypanosoma brucei rhodiense*	*Trypanosoma cruzi*	~21 *Leishmania* spp.
Vector	Tse-tse fly (~20 Glossina spp.)	Reduviid insects	Phebotomine sandflies
Geographical distribution	Sub-saharan Africa (~20 countries)	South America (~19 countries)	South and Central America, Europe, Africa and Asia
Number of infected persons	80,000	8 Million	12 Million
Annual deaths	30,000	14,000	51,000

(Data from www.who.int/neglected_diseases/diseases/en/).

The life-cycle of *Trypanosoma brucei* is complex. It involves at least two main developmental stages, the insect-infecting, or procyclic form and the mammal-infective, or bloodstream form. There are multiple morphologic and metabolic differences associated with the distinct developmental forms (7,8). For example, the procyclic form derives most of its ATP through oxidative phosphorylation, while this pathway is inactive in bloodstream forms, which heavily rely on a specialized organelle, the glycosome, to replenish their pool of ATP through substrate-level phosphorylation events in glycosysis (9). Both developmental forms of the parasite can be cultivated *in vitro* using cell-free, semi-defined media, enabling their study in different stages of the

life-cycle, however because the procyclic form grows to higher densities, most biochemical work has historically focused on this stage.

Apart from their medical relevance, trypanosomes have proved of considerable intrinsic scientific interest for their evolutionary history, quirky metabolism, complex life-cycles and unique mechanisms of gene expression, such as RNA editing (9-11).

RNA Editing

The majority of eukaryotic RNAs are transcribed as precursors that must undergo a variety of post-transcriptional processing events (such as splicing, capping, polyadenylation, etc.) that are required to convert them into mature RNA species. The term "RNA Editing" was originally introduced in the mid-1980s to describe the insertion and deletion of uridine residues into mitochondrial messenger RNAs (mRNAs) of kinetoplastid protozoa. It soon became clear that RNA editing is a widespread biological phenomenon; examples of editing from bacteria to humans have been reported for mRNAs, transfer RNAs (tRNAs), ribosomal RNAs (rRNAs), and micro- and short-interfering RNAs (miRNAs and siRNAs). Thus, the use of the term "RNA Editing" has extended to describe a multitude of processes resulting in the formation of RNAs with sequences different from those coded by their DNA templates. In general, RNA editing can be thought of as any post-transcriptional process that modifies the chemical composition of RNA molecules and results in a change of their encoded information and/or function (9,10,12,13).

These processes, which are largely unrelated and occur by a variety of mechanisms, can alter the function or coding potential of the majority of cellular RNAs and can be divided in three broad categories: base deaminations, nucleotide modification and nucleotide insertion/deletion RNA Editing.

Functions of RNA Editing

Editing of Messenger RNA (mRNA)

Most RNA editing occurs to sequences within mRNAs, which results in the production of altered protein products. Entire open reading frames can be created by extensive nucleotide insertions in mitochondrial transcripts of trypanosomes and Physarum (10). Paramyxoviruses exploit editing of mRNA to direct frameshifting between alternative ORFs (14,15). Start and stop codons are established by uridine insertion/deletion in trypanosomes and cytidine to uridine (C-to-U) changes in plant organelles and humans (9,16,17). U-to-C substitution can remove stop codons in plants (17-19). Altered amino acid coding potential is exploited by editing; amino acid substitutions occur due to C-to-U, U-to-C, and adenosine to inosine (A-to-I) changes (13,16,17,20-22), additionally, splice sites may be altered in response to A-to-I conversion (23). Editing within 5'-and 3'-untranslated regions (UTRs) may affect mRNA localization, stability, translation competence and /or processing (13,23,24).

Editing of Transfer RNA (tRNA)

The existence of base modifications in tRNA has been long recognized, but examples of sequence changes in tRNA involving only canonical nucleotides have been

reported in organisms ranging from *Acanthamoeba* to fungi, land plants, metazoans and potentially bacteria (10,17,25-32). Editing creates essential secondary and tertiary structural elements, involving both loop nucleotides and base-paired stems (10,26) and can occur by nucleotide insertion/deletion (25,26,28,29), or base conversion (10,26,27,33,34). Some editing events alter tRNA recognition by aminoacyl tRNA synthetases, whereas others affect 5'-and 3'-processing or subsequent base modification (10,26).

Editing of Ribosomal RNA (rRNA)

Ribosomal RNA editing is less common than editing of mRNAs and tRNAs. There is a single C-to-U substitution in the highly conserved 530 loop, a region critical in tRNA selection and proofreading in the mitochondrial small subunit rRNA of *Dictyostelium* (10). Both the large and small mitochondrial rRNAs of *Physarum* contain extensive nucleotide insertions (35,36). The edited regions participate in the formation of conserved helices, the 530 loop, and an AA dinucleotide involved in binding tRNA at the A site (10).

Editing of MicroRNAs and siRNAs (miRNAs and siRNAs)

MicroRNAs are major targets for adenosine deaminases acting on RNA (ADARs) (20,23,24,33,37). Editing of miRNAs can affect their biogenesis by blocking the action of the miRNA processing machinery as well as alter their specificity and targets through nucleotide changes (20,21,24,37). Approximately 6% of all mature miRNAs in humans are edited. An analysis of miRNA precursors show that up to 50% of all primary-miRNAs (pri-miRNAs) contain A-to-I modifications (23). Editing

appears to be miRNA-specific; some pri-miRNAs are extensively edited, such as pri-miR-223, pri-miR-1-1 and pri-miR-143, while editing is not observed in other precursors, such as pri-miR-181 and pri-miR-122 (37). Editing sites residing in the miRNA "seed" region can affect miRNA target recognition and function (21,38). The edited and unedited versions of miR376a-5p interact with and regulate the expression of a unique set of mRNA targets (38).

A-to-I editing of siRNAs can modulate their RNAi efficacy (13,22). Studies with Drosophila extracts suggest that edited siRNAs are less effective at triggering an RNAi response (39). If dsRNA is incubated with an ADAR prior to adding it to a Drosophila extract, siRNAs are not detected. It is thought that if ADARs deaminate a dsRNA before entering the RNAi pathway, proteins that bind dsRNA, such as Dicer, cannot bind effectively and gene silencing does not occur (20).

U-Insertion/Deletion RNA Editing

RNA editing was first described over two decades ago in mitochondrial transcripts of *Crithidia*, a kinetoplastid organism. Benne and co-workers reported the presence of four uridines in the mRNA of cytochrome oxidase II that were not encoded in the gene (40). Following this, another group found extensive editing of cytochrome oxidase III mRNA that resulted in the insertion of 547 uridines and the deletion of 41 uridine residues (41).

The trypanosome mitochondrial genome encodes nine mRNAs that are extensively edited, three mRNAs that are edited at a limited number of sites, and six mRNAs that are never-edited (42). Editing is essential in both insect and bloodstream developmental stages of the parasite, and results in the formation of initiation codons and extended open reading frames in mRNA, necessary for essential mitochondrial processes such as electron transport and oxidative phosphorylation (9,12,42,43). In essence, RNA editing produces translatable mRNAs to allow the synthesis of mitochondrial proteins. Another potential *raison d'être* of RNA editing might be to expand the coding capacity of the mitochondrial genome by producing alternative proteins from a single gene (9,13,44). For example, alternatively edited protein 1 (AEP-1), a protein that functions in mitochondrial DNA maintenance, is encoded by an isoform of the cytochrome c oxidase (COIII) mRNA consisting of pre-edited and edited sequence (44).

Although the evolutionary origin of this process is still a mystery, RNA editing may be a mechanism to prevent the loss of genes that are only necessary during certain parts of the life-cycle by their segmentation into multiple fragments, all of which would be required for expression (45).

Mechanism of U-Insertion/Deletion

RNA editing in kinetoplastids is directed by small, trans-acting, guide RNA (gRNA) molecules that, through canonical Watson-Crick and G-U base-pairing with pre-mRNAs, specify the site, type and extent of editing by directing the enzymatic activities residing in the RNA editing core complex (RECC), a multi-protein, multi-enzyme complex of approximately one megadalton (46,47). The model of trypanosome RNA editing involves an "enzymatic cascade" (Fig. 1) (42,48) and begins by the formation of an "anchor" duplex between the 5' residues of gRNA and complementary residues in the pre-mRNA downstream of the editing site (ES), which is defined by helical irregularities in the pre-mRNA:gRNA complex (48-51). A complete editing cycle involves three enzymatic steps: 1) cleavage of the pre-mRNA at the scissile phosphodiester by an RNase III endoribonuclease, 2) addition of uridylate to the 3' end of the upstream cleavage fragment by a terminal uridylyl transferase (TUTase) using UTP as substrate, resulting in an insertion, or removal of UMP from the 3' end of the upstream cleaved fragment by a 3'-5' exouridylylase to result in a deletion, and finally, 3) ligation of the fragments by an RNA ligase. The mature message is created by successive rounds of editing, often involving numerous sites and gRNAs, and a general 3'-5' polarity.

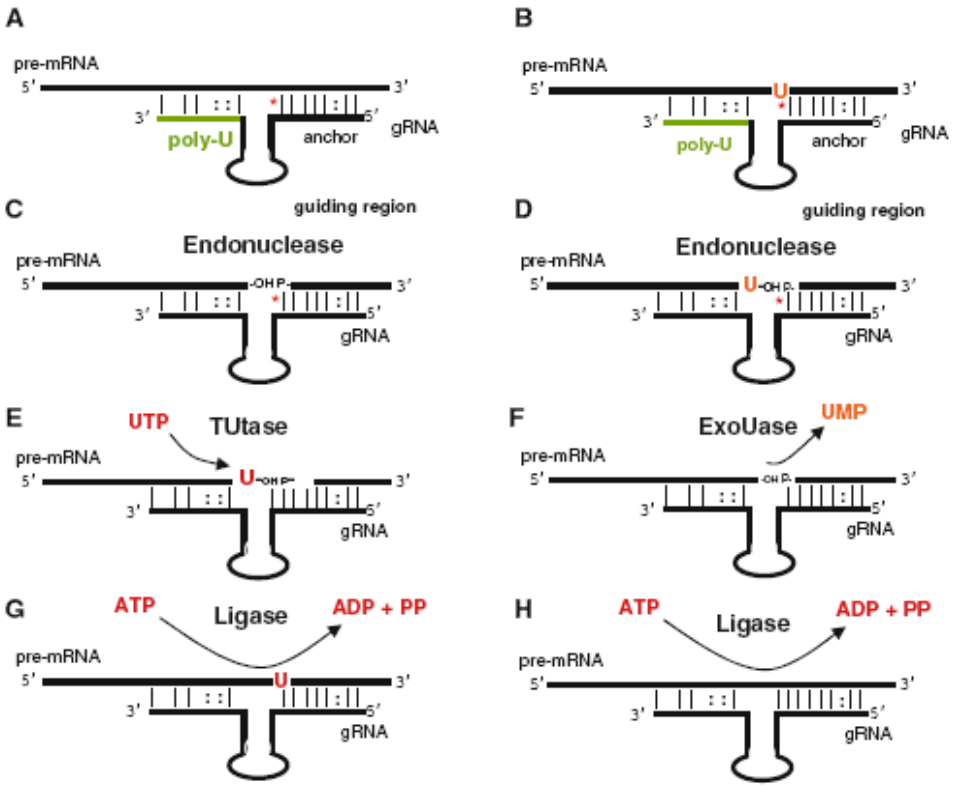

Figure 1. Mechanism of U-insertion/deletion RNA editing. A, B. gRNAs anneal to their cognate pre-mRNAs via a short anchor region on the gRNA 5' end. The 3' poly-U tail of gRNA stabilizes the gRNA:mRNA interaction. C, D The first mismatch (*) of the anchor duplex is the signal for endonuclease cleavage of the pre-mRNA. E, F After cleavage, U's are added or deleted from the 3' end of the upstream cleavage product by a TUTase or exonuclease. G, H The 5' and 3' mRNA fragments are sealed by an RNA ligase. (From Ochsenreiter and Hajduk, 2008).

The RNA Editing Core Complex (RECC)

U-insertion/deletion RNA editing is catalyzed by the RNA editing core complex (RECC), a multi-subunit protein complex composed of approximately 20 proteins which sediments at ~20S in glycerol gradients (Table 2) (52-55). The composition of RECC has been revealed by mass spectrometric studies of RECC purified by a variety of conventional chromatographic methods, immunoprecipitation using monoclonal antibodies or tandem affinity purification. Many RECC subunits have domains indicative of mediating interaction with nucleic acids or protein-protein interactions. Genetic investigations have shown that with few exceptions, the majority of RECC subunits are necessary for RECC structural integrity and essential for parasite viability (42).

Table 2. Components of RNA Editing Core Complexes (RECC).

Enzymatic RECC Subunits	Name	Former name	Function	Motif
	REN1	MP90, KREPB1	Deletion endonuclease	RNase III, dsRBD, U1-like
	REN2	MP61, LC-6a, KREPB3	Insertion endonuclease	RNase III, dsRBD, U1-like
	REN3	MP67	cis-guide endonuclease	RNase III, dsRBD, U1-like
	REX1	MP100, LC-2	ExoUase	5′3′ exo, endo/ exo/phos
	REX2	MP99, LC-3, band I	ExoUase	5′3′ exo, endo/ exo/phos
	RET2	MP57, LC-6b	TUTase	NT, PAP-core, PAP-assoc
	REL1	MP52, LC-7a, band IV	Ligase	Ligase, tau, K
	REL2	MP48, LC-9, band V	Ligase	Ligase, tau, K
Non-enzymatic RECC Subunits	MP81	KREPA1, LC-1, band II	Interaction	OB-fold, zinc-finger
	MP63	KREPA2, LC-4, band I	Interaction	OB-fold, zinc-finger
	MP42	KREPA3, LC-7b, band VI	Interaction	OB-fold, zinc-finger
	MP24	KREPA4, LC-10	Interaction	OB-fold
	MP19	KREPA5	Interaction	OB-fold
	MP18	KREPA6, LC-11, band VII	Interaction	OB-fold
	MP46	KREPB4, LC-5	Interaction	RNase III?, Pumilio, U1-like
	MP44	KREPB5 LC-8	Interaction	RNase III?, Pumilio, U1-like
	MP49	KREPB6, LC-7c	Interaction	U1-like
	MP47	KREPB8	Interaction	U1-like
	MP41	KREPB9	Interaction	U1-like

[The names used reflect a recent proposal to unify the various systems of nomenclature in the literature (56). The function of a majority of proteins has been experimentally determined, but in some cases, it is based on predictions based on sequence motifs. The function "interaction" refers to RNA- and protein-protein interactions. OB-fold, oligonucleotide/oligosaccharide-binding fold; zinc-finger, C_2H_2-type zinc finger domain; RNase III, endoribonuclease III domain; dsRBD, double-stranded RNA binding domain, U1-like, U1-like zinc finger motif; Pumilio, Pumilio RNA binding motif; ligase, ligase signature motif; tau and K, putative microtubule-associated tau and kinesin light chain domains; 5'3'exo, 5'-3' exoribonuclease motif; endo/exo/phos, endonuclease/exonuclease/phosphatase domain; NT, nucleotidyltransferase motif; PAP-core and PAP-assoc, poly(A) polymerase core and associated domains (54,56).]

Catalytic RECC Components

Endoribonucleases

The first enzymatic step in RNA editing involves the gRNA-directed endoribonucleolytic cleavage of the pre-mRNA at the ES by RNase III-like RECC proteins that recognize and cleave deletion, insertion or cis-guided sites (Fig. 1A) (42). RNA editing endonuclease (REN) 1, REN2 and REN3 are the editing endonucleases for deletion, insertion and cis-guided sites, respectively (57-59). These proteins share a similar domain architecture, with an N-terminal U1-like Zinc-finger domain, a central RNase III domain that is essential for catalysis and a C-terminal double-stranded RNA binding domain (dsRBD) (Fig. 2) (54,60).

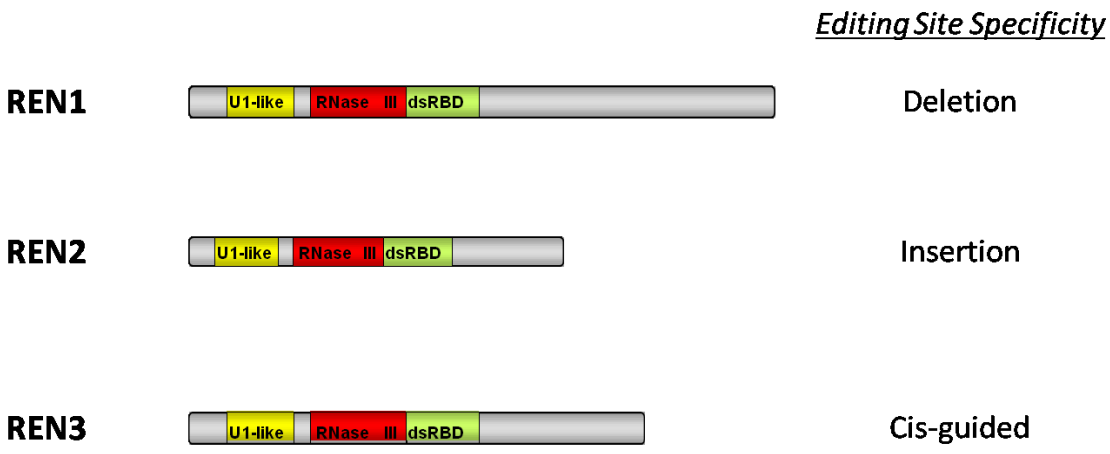

Figure 2. Domain structure and editing site specificity of REN proteins. Cleavage specificity is indicated to the left of the protein diagram. Domains are indicated in colored boxes: yellow, U1-like zinc-finger; red, RNase III nuclease domain; green, double-stranded RNA binding domain.

Expression of REN proteins is essential. RNA interference (RNAi) of REN1 or REN2 expression leads to the loss of edited mRNAs (58,59). Bloodstream parasites in which the REN3 genes are knocked-out show inhibited growth and defects in COII editing (the sole mRNA reported so far that requires a cis-guided editing event) (57).

Affinity purification of RECC by TAP-tagged REN proteins results in the exclusion of the two other RENs but an otherwise similar RECC subunit composition (61). The biological relevance of the compositional heterogeneity of TAP-tagged REN RECCs is unknown, but it has been proposed that REN proteins may "shuttle" in and out of the core RECC assembly during the course of editing as the need arises to cleave a specific type of editing site (52,61).

Exoribonucleases

A U-specific exonuclease activity removes unpaired U's in the mRNA after cleavage of a deletion editing site. Two RECC proteins, RNA editing exonuclease (REX) 1 and REX2 contain N-terminal 5'-3' exonuclease and C-terminal endo/exo/phosphatase motifs (54). Although both proteins show exonuclease activity *in vitro* (62,63), it is currently thought that REX1 plays the major role in U-removal *in vivo*, as repression of its expression results a growth phenotype and in a decrease of U-removal activity *in vitro* by cell extracts, whereas down-regulation of REX2 expression by RNAi does not inhibit cell growth (52,64).

Terminal Uridylyl Transferase

Following cleavage of an insertion site, the RECC enzyme RNA editing uridylyltransferase (RET) 2 catalyzes the gRNA-encoded addition of uridylates to the 3' end of the upstream cleavage fragments (65). This enzyme is related to RET1, an enzyme responsible for 3'-end uridylylation of mitochondrial transcripts (66) and both proteins share nucleotidyl transferase and poly-(A) polymerase domains (54). RNAi of RET2 expression results in growth inhibition, a reduction in edited RNA levels and loss of pre-cleaved insertion editing activity *in vitro* (54,65). A recombinant version of RET2 adds U's to pre-cleaved editing substrates in a gRNA-directed manner and a single U to single-stranded RNA. Interaction with MP81, a RET2-binding partner *in vivo*, stimulates RET2 activity *in vitro* (65) which is consistent with the accumulation of subcomplexes upon the over-expression of REL2 that are active in pre-cleaved insertion and RET2, MP81 and REL2 (67).

RNA Ligases

The final step in the enzymatic editing cycle involves the ligation of the mRNA fragments after U processing. RECC contains two closely related ligases, RNA editing ligase (REL1) and REL2. These proteins contain signature ligase motifs (68), share 40% amino acid identity and are related to RNA ligase 2 of phage T4 (54). REL proteins however, do not have an oligonucleotide/oligosaccharide binding (OB)-fold commonly found in other ligases. It is thought that their binding partners, MP81 and MP63, which possess OB-folds, supply this domain *in trans* to the RELs (67).

Functionally, REL1 has been linked with U-deletion and REL2 with U-insertion (63,69-74) and their over-expression promotes the formation of separate subcomplexes active for deletion or insertion (67). Expression of REL1 is essential for parasite survival and editing *in vivo*. KREL2 is not essential, suggesting that KREL1 can complement for the loss of KREL2, but not vice versa. In line with this, recombinant REL1 and REL2 show differences in ligation activity of substrates containing gaps and overhangs (75). In addition, in the context of RECC, they have different ATP concentration optima as well as show different inhibition by pyrophosphate *in vitro* (73).

Non-catalytic RECC Components

RECC components that do not contain motifs suggesting enzymatic activity nevertheless possess motifs known to mediate interactions with RNA- and/or protein. Many of these proteins are important for RECC integrity and, possibly, for its assembly (42,52,54). Although their nomenclature has been a contentious issue (56), at present it is useful to use a classification system based on sequence similarities that divides these proteins in two groups: the kinetoplastid RNA editing proteins (KREP) A and B families (54).

KREPA Family

Members of the A family (KREPA 1-6, alternatively mitochondrial protein (MP) 81, MP63, MP42, MP24 and MP18) contain a C-terminal OB-fold, suggesting their interaction with RNA (76). MP81, MP63 and MP42 also contain two N-terminal C_2H_2 zinc-finger domains which seem to be important for the editing process; mutations in the zinc-finger domains of MP63 affect RECC stability while mutations in the zinc-finger domains of MP42 affect the progression of editing (77,78)

KREPB Family

Members of the KREPB family (KREPB4-8, or MP41, MP44, MP46, MP74 and MP49) contain an N-terminal U1-like Zinc-finger domain, which suggest that they may interact with RNA and/or protein (52,54). MP44 and MP46 contain a C-terminal Pumilio RNA binding motif and an RNase III domain that is thought to be enzymatically inactive, as amino acid residues critical for catalysis are mutated. Both proteins play a major role in RECC stability; knockdown of their expression by RNAi leads to the disappearance of the RECC and its subunits (79,80). Apart from the U1-like Zinc finger domain, MP41, MP47 and MP49 lack discernible amino acid motifs. These proteins are thought to associate preferentially with affinity-purified RENs. MP41 is only found in TAP-REN1 RECCs, MP47 is found in TAP-REN2 RECCs and MP49 is found only in TAP-REN3 RECCs (52,57,61).

RNA Editing In Vitro

Mechanistic studies revealed that U-insertion and U-deletion involve different enzymatic activities (48,73,74,81-84). Specifically, the endonuclease activities in the two editing types exhibit different behavior in the presence of adenosine nucleotides, the U-addition and U-removal steps are not reverse reactions catalyzed by the same enzyme and separate ligases are used for joining insertion and deletion substrates (62-65,67,70,72,73,75,84-87).

U-insertion and U-deletion RNA editing activities can be assayed *in vitro* using synthetic mRNA and gRNA substrates in the presence of mitochondrial extract or purified RECC and ATP (in addition to UTP for insertion, and ADP for deletion) (85,88). There are two basic variations of the editing assay *in vitro*: "full-round" and "pre-cleaved" assays (Fig. 3). In addition, individual activities (endonuclease, exonuclease/TUTase, or ligase) can be assayed by modification of these standard assays (85,88,89).

Full-Round Editing Assay

Full-round RNA editing assays *in vitro* are relatively inefficient as they asses the action of multiple enzytic activities simultaneously (cleavage of mRNA, U-insertion or deletion and ligation). This assay has been limited to editing of a single site directed by a single gRNA (85,88) until recently, where editing of two sequential editing cycles was reported (90). Although several mRNA substrates and their cognate gRNA have been used (91), the most efficient and widely used substrates are derived from the ATP synthase subunit 6 or A6 mRNA (49-51).

Cleavage Assay

The standard full-round editing assay can be modified to examine endoribonuclease cleavage activity. Adenosine diphosphate (ADP) stimulates the cleavage of deletion substrates while partially inhibiting insertion cleavage (51,92). In addition, RNA ligases (which may be pre-charged in RECC preparations) are deadenylated by a pre-incubation of RECC preparations with high concentrations of pyrophosphate prior to the assay (92).

Pre-cleaved Editing Assay

The pre-cleaved assay bypasses the endonucleolytic cleavage step, which tends to be the limiting step for editing *in vitro* (85), by supplying the substrate RNA as three fragments that mimic a cleaved substrate (83,84,93). This assay scores either U-addition or deletion and RNA ligation only and tends to be more efficient than the full-round assay. In addition, the omission of ATP from the reaction mixture leads to a decrease in ligase activity, making possible the examination of exonuclease or TUTase activities individually.

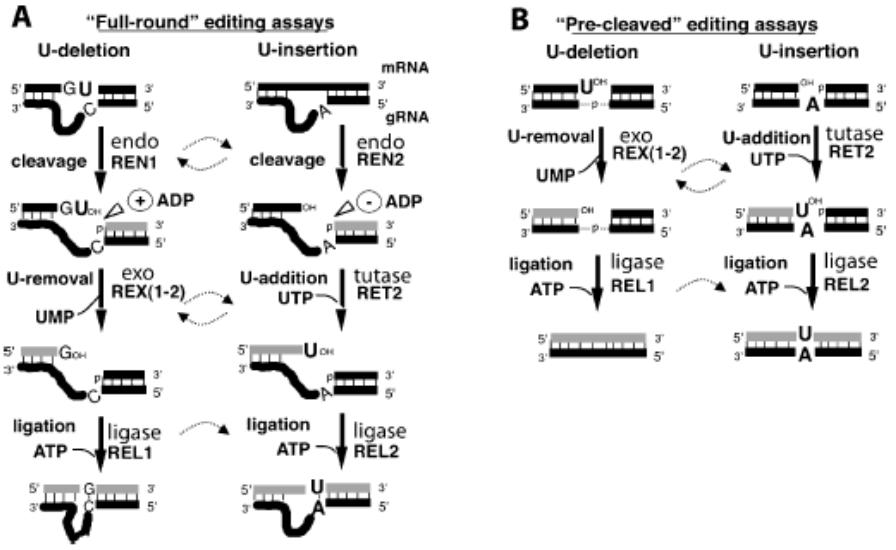

Figure 3. Full-round and pre-cleaved editing assays. (A) A full-round of editing is the product of a process composed of three gRNA-directed enzymatic steps: pre-mRNA cleavage, U-insertion/addition and ligation. Enzymes responsible for individual steps are shown. Arrowhead points to the editing site. Nucleotide co-factors are indicated; +:stimulation, -:inhibition. (B) Pre-cleaved assays score only the U-removal/addition and ligation steps. [From Hernandez and Cruz-Reyes, 2008 (42)].

RNA-Protein Interactions

With the goal of understanding the mechanism of substrate recognition by RECC, the Cruz-Reyes laboratory has pioneered the study of protein-RNA interactions in the context of assembled editing complexes. We have identified four major RNA-protein interactions using purified RECC and synthetic RNA substrates with model editing sites [Chapter II; (94)]. The cumulative data suggest that initial events in substrate recognition and binding are sequence-independent, follow relatively simple RNA secondary structural determinants and are mediated in large part through RECC RNase III proteins [Chapters III and IV; (92,95)].

Regulation of RECC Enzymatic Activities-The Importance of cis-Acting Elements

RECC enzymes are directed and regulated through a combination of the structure of the mRNA:gRNA hybrid, the architecture or the complex and conformational changes (in the substrate and RECC) brought about by binding and/or catalysis. It has been proposed that the RECC is divided into deletion and insertion "halves" (67,81,82). The structure of the mRNA:gRNA hybrid may serve as the initial "landing pad" for RECC via binding to double-stranded RNA region. The identity of the editing site is then established by proteins that bind to the unpaired region in the hybrid and direct the substrate to its appropriate RECC "half" or subcomplex before mRNA cleavage occurs.

Editing endonucleases target phosphodiesters at single/double-stranded junctions (48-50,92). A combination of substrate determinants and/or endonuclease-binding partners may impart specificity to RECC endonucleases for their particular substrates. *In vitro*, high concentrations of adenosine diphosphate (ADP) are required for cleavage of deletion substrates but partially inhibit the insertion cleavage activity (81). In this regard, ADP may promote an editosome conformation that activates deletion cleavage or inactivates insertion cleavage by an allosteric mechanism. In addition, recent evidence suggests that a number or RNA helicases may be involved in RNA editing (and co-purify with RECC; discussed in Chapter V). An intriguing possibility is that the effects observed upon addition of ADP may be a consequence of the regulatory role of helicases responsive to ATP or the ATP/ADP ratio. These helicases may exert a regulatory role by modifying the architecture of the RNP, inducing conformational changes in the mRNA:gRNA complex as well as in the RECC that may be centrally involved in

controlling binding and/or the progress of catalytic events. The modulation of cleavage by ADP also raises an interesting possibility that RNA editing may be regulated by the cellular status through the sensing of key metabolites and/or their ratios.

After mRNA-cleavage, determinants in the mRNA/gRNA substrate play a pivotal role in the regulation of editing enzymes. For insertion, the number of unpaired adenylate or guanylate (since G-U pairs are also found) residues on the gRNA strand dictates the number of uridylates added to the mRNA strand by the insertion terminal uridylyl transferase, while for deletion, unpaired uridylate residues on the mRNA strand are removed by the deletion exonuclease. Thus the internal loop, an unpaired region between an mRNA and a gRNA, is solely responsible in specifying the type and extent of editing. The interconversion of editing sites by alteration of features in the internal loop supports this model (91).

In addition to specifying the mRNA cleavage site, the gRNA strand is of utmost importance for the accuracy and fidelity of editing. The gRNA in the tri-partite cleavage product directs the addition or deletion of uridylates to the 5' mRNA fragment. After cleavage, RECC enzymes require a way of "counting" how many U's must be modified in order to accurately edit a site. This task is accomplished by a combination RNA structure and the specificity of the enzymes involved. For example, the recombinant version of the RET2 adds a single U to single-stranded RNA, but when a construct mimicking a tri-partite cleavage product with a gRNA containing two unpaired A's was used as substrate, the enzyme added the gRNA-directed number of Us (96). The

formation of a nicked duplex by the base-pairing of the added U's in the mRNA and A's in the gRNA may.

A recombinant version of the deletion 3'-5' exouridylylase REX1 was found to be single-strand specific, and that U's in nicked duplexes were not cleaved (62). Thus a common strategy for the regulation of the uridylate addition/removal steps in RNA editing may be through the inhibitory effect of a duplex that does not allow access to the mRNA 5' fragment and which becomes a substrate for the subsequent ligation step. Some mis-editing occurs, *in vivo* and *in vitro* (9). Excess U's added during an insertion cycle can be removed by the deletion exonuclease (97), and one can imagine a scenario where the insertion TUTase can reinstate extra U's removed by the deletion exonuclease. The action of deletion enzymes in the insertion cycle and vice-versa may constitute a form of "proof-reading" and quality control step where aberrant products of one pathway may be recognized and correctly "re-edited" by the other pathway.

After the U-processing steps, the tri-partite mRNA:gRNA complex becomes a substrate for RNA ligases. Genetic evidence suggests that REL1 is essential and may be the main RNA ligase in RNA editing (68,72). Although both REL1 and REL2 are active in recombinant form and show similar specificity, only REL1 in the presence of REX1, ligated a substrate with overhanging U's (63,69). This suggests that either the mRNA:gRNA structure is made, by binding to REX1, more amenable to ligation by REL1 or that REL1 and REX1 interact and their complex can bind and ligate the substrate with overhanging U's while REL1 in isolation cannot. This implies that the catalytic activities of editosome enzymes may depend not only in the intrinsic properties

of the enzymes, but also on the binding of specific protein partners that form a more complex active site and/or substrate binding site. It will be highly interesting to use the emerging picture of protein-protein interactions among RECC subunits (98) and establish if their activities can be modulated, enhanced or regulated by such binding events.

RNA Editing Accessory Factors

Several RNA editing accessory factors have been identified (Table 3); however, there is little evidence to suggest that they are essential for RNA editing. Instead, they are thought to play a role in RNA stability, the formation of mRNA:gRNA hybrids and/or the assembly of substrates onto the RECC (Table 2).

Table 3. Accessory Proteins in RNA Editing.

Name	**Function**	**Motif(s)**
MRP1/2	mRNA-gRNA matchmaking	Lysine-rich
RBP16	Cyb mRNA stability/ RNA-RNA Annealing?	Cold-shock
REAP1	mRNA stability	N/A
mHEL61/REH1	Editing progression?	DEAD-box

Their function and apparent amino acid motifs are indicated.

One of the first such factors found was the gRNA-binding protein MRP1 that, together with MRP2, form the MRP (mitochondrial RNA-binding protein) complex (99). The MRP complex facilitates RNA-RNA annealing *in vitro*, and it is thought to aid in the matchmaking of mRNA-gRNA pairs. Bloodstream *T. brucei* with an MRP1 deletion are viable but they show reduced levels of edited RNA and cannot progress to the insect form of the organism (54). RNAi knockdown of MRP2 expression inhibits cell growth and affects the abundance of edited and unedited versions of mitochondrial RNAs (54,100). The MRP complex may exert a regulatory role in RNA editing by binding pre-mRNA and gRNA transcripts and inducing their association to form an editing substrate either directly, by bringing the editosome and the mRNA:gRNA substrate together, or passively, by the formation of the mRNA:gRNA hybrid that becomes a substrate for the editosome. Furthermore, the formation of a double-stranded mRNA:gRNA hybrid may enhance the stability of transcripts in the mitochondrial matrix by directing them to the editing pathway and avoiding RNA decay/turnover pathways (101). In addition, by binding mRNA and/or gRNAs, irrespective of their annealing, the MRP complex can control the timing of editing for a particular transcript.

RBP16 contains a cold-shock domain and has affinity for oligo (U)-containing RNA (102,103). This protein has not been shown to directly bind RECC, but *in vitro*, it stimulates the editing reaction (104). RNAi repression of RBP16 leads to changes in the ratio of edited and pre-edited forms of the transcript Cyb in addition to a reduction in the levels of the never-edited mitochondrial mRNAs; however, gRNA levels are not affected (102,103). Recently, it was shown that RBP16 is methylated at specific arginine

residues, this affects its RNA binding properties and may facilitate RNA-RNA annealing (105,106). Unlike the MRP complex, RBP16 appears to function at stages both upstream and downstream of the editing pathway, possibly by binding and stabilizing a specific subset of Cyb transcripts before and after editing. Like the MRP complex, it may exert a regulatory role by binding and stablilizing mRNA or gRNA, thereby controlling the timing of their entry into the editing pathway.

RNA editing-associated protein 1 (REAP1) is a poorly-understood RNA-binding protein first proposed to play a role in RNA editing over a decade ago (107). REAP1 knockouts are viable but display a growth phenotype and show an altered profile of mitochondrial RNAs, suggesting that REAP1 functions in mitochondrial RNA stability (108).

A suite of mitochondrial RNA-binding multi-protein complexes of overlapping composition and function have been reported by several labs, including our own (see Chapter V). These complexes seem to be intimately involved in the metabolism of gRNA (109-112). We propose these complexes reflect a dynamic network of mitochondrial complexes involving RNA maturation, stability and translation (see Chapter V).

Despite great progress in the elucidation of the composition and enzymatic activities of RECC, the dynamic character of RNA editing is obscure. It is likely that regulation of RNA editing is achieved through the regulation of the various steps in the editing reaction. Both substrate RNA molecules and RECC subunits possibly undergo conformational changes during annealing, binding, as well as during progression of the

reaction, including the removal of guide RNA after editing is complete. RNA helicases may facilitate these remodeling events.

DExH/D-box proteins are typically viewed as ATP-dependent RNA helicases, but recently it has been recognized that they may function as RNP remodeling factors (113). A mitochondrial DEAD-box protein, mHEL61p, has been proposed to be involved in RNA editing (114). Trypanosomes deleted in mHEL61p show strongly reduced levels of edited mRNAs, while RNAs that unedited RNAs are not affected (114). It is possible that other RNA helicases/RNPases may function in RNA editing. Our laboratory recently found evidence for the association of RECC with a mitochondrial DExH-box protein containing a C-terminal dsRNA-binding domain (Chapter V) and the existence of an additional DExH-box helicase bearing a dsRBD. Thus it is possible that multiple RNA helicases/RNPases may function at different steps in RNA editing to facilitate RNP conformational changes that lead to a certain binding or catalytic event. It will be of great interest to determine the steps along the editing reaction where RNA structure or RNP composition are altered by the action of one or more RNA helicases and how this regulates RNA editing.

Editing in Different Trypanosome Life-Stages

T. brucei shows significant regulation of most mitochondrial transcripts during its life cycle. Since most edited sequences encode subunits of respiratory chain complexes, it is thought that RNA editing helps regulate energy metabolism in this and related organisms (2,7,8).

In the bloodstream stage, mRNAs encoding components of NADH dehydrogenase are highly edited, suggesting a role for this respiratory chain component in this life-stage although it lacks messages for apocytochrome b (Cyb) and cytochrome oxidase (COII). One edited mRNA in trypanosomes encodes a ribosomal subunit (RPS12), implying that RNA editing may also serve a role in the control of mitochondrial translation. Messages encoding apocytochrome b (Cyb) and a subunit of cytochrome oxidase (COII) are edited by insertion in the insect-borne, or procyclic, stage only. Other messages, such as those encoding for ATP synthase subunit 6 (A6), cytochrome oxidase (COIII) are edited in both life-stages.

The molecular basis for the developmental control of RNA editing is unknown. There is evidence to suggest that editosomes may differ in composition and activity between the two life-stages (115). Also, although the steady-state levels of several edited mitochondrial mRNAs vary significantly between the procyclic and 27mmune27nscr life-stages, the abundance of gRNA species is similar (116). Developmental differences in RNA editing thus may be due to changes in usage of gRNAs, RNA turnover and/or processing of precursor RNAs. A change in substrate specificity and activity of the RECC brought about by changes in its protein composition may partially account for the observed differences, but unless gross changes in specificity, affinity and activity of editosomes occur, it is highly likely that factors acting upstream of RECC involved in RNA stability, formation of mRNA:gRNA hybrids or "presentation" of substrates to the RECC might be the "master regulators" in RNA

editing in the sense that they would determine the cohort of RNAs that reach the editosome or that are targeted to turnover.

Other Types of RNA Editing

Table 4 summarizes the major types of RNA editing.

Table 4. Types of RNA Editing

Type of RNA Editing	Mechanism	Organism
U-Insertion/Deletion	Enzymatic cascade; Cleavage-ligation	Trypanosomes
C-to-U	Hydrolytic base deamination	Mammals; Plant organelles
U-to-C	"	Lower plants
A-to-I	"	Higher eukaryotes
Viral Polymerase "Stuttering"	Pausing and reiterative polymerization by viral RdRp	Paramyxoviruses; Ebola virus
C-Insertion and Dinucleotide Insertion	Co-transcriptional	*Physarum* molds
tRNA Nucleotide Deletion/Insertion	Nuclease, nucleotidyltransferase	*Acanthamoeba*

The mechanism of editing and the organisms in which it occurs are indicated.

C-to-U Editing

C-to-U Editing in Mammals

The best understood C-to-U RNA editing event involves the mammalian nuclear transcript encoding intestinal apolipoprotein B (apoB) (10,16). Editing of ApoB mRNA alters a CAA to a UAA stop codon, generating a truncated protein, apoB48. Editing of ApoB mRNA has important effects in lipoprotein metabolism and intestinal and hepatic lipid transport in mammals (16). C-to-U editing of apoB mRNA requires a single-

stranded RNA region as well as proteins that assemble into a functional editing complex called the editosome. This complex is minimally composed of the catalytic deaminase, apobec-1, and an adaptor protein, called apobec-1 complementation factor (ACF), which binds both the deaminase and the mRNA substrate.

Other examples of C-to-U RNA editing in mammals has been found and are thought to share the same machinery as apoB RNA editing. In patients with neurofibromatosis type 1 (NF1), a CGA codon is deaminated to a UGA stop codon in the mRNA encoding the protein neurofibromin. It is thought that this mRNA encodes a truncation product that lacks a critical GTPase activation domain; however, there is no evidence that the truncated protein is produced (16). Another target for C-to-U editing, NAT1, was found when apobec-1 was overexpressed in mouse and rabbit hepatocytes (117). NAT is homologous to the translational repressor eIF4G, and editing of this mRNA creates stop codons that reduce protein abundance (16).

C-to-U Editing in Plants

C-to-U conversions have also been observed in mitochondrial and chloroplast mRNAs of higher plants (17,18). Although the enzymes involved have not been identified, C-to-U editing in plants is thought to involve a cytidine deaminase (10,17). Editing appears more prevalent in mitochondrial transcripts than in chloroplast transcripts: 500-1,000 C-to-U editing events are estimated to occur in mitochondria but only 4-25 editing events have been reported in chloroplasts (10,17,18). Comparisons of editing sites between various plants species suggest that C-to-U modifications are

conserved (17). Most substitutions are in the first or second positions of codons and appear to increase the hydrophobicity of the encoded proteins (13).

U-to-C Editing

Lower plants show a large number of uridine-to-cytidine (U-to-C) RNA editing events in both mitochondrial and chloroplast transcripts (17,18). This phenomenon has also been described for certain transcripts encoding neuropeptides in frogs as well as Wilm's tumor susceptibility gene (WT1) in mammals (10). The mechanism responsible for U-to-C editing is unknown but it is speculated that it may be related to the reaction catalyzed by CTP synthetase, an enzyme of pyrimidine metabolism that produces CTP by transferring an amino group from glutamine to UTP.

A-to-I Editing

Adenosine-to-Inosine (A-to-I) substitution RNA editing may be the most prevalent type of editing in higher eukaryotes and, especially, in the nervous system of mammals (20,22-24,33,34,118). Pre-mRNA exons, repetitive sequence elements in untranslated regions and miRNA precursor transcripts are three major substrates for A-to-I editing (13,24).

The extent of editing can vary in response to both the secondary structure of the RNA and the length of duplexed regions; short duplexes (15-40 base-pairs) are edited at a few sites, while duplexes greater than 50 base-pairs are edited extensively (24,118). RNA regions that form perfectly base-paired duplexes are more extensively edited, while more complex structural regions may be edited with higher selectivity (13,20).

Editing of mRNA coding regions alters its encoded information because inosine is read as guanosine (G) during splicing and translation. In addition, editing of non-coding regions appears to have a large influence in various processes, including the cellular localization, translation efficiency and stability of RNAs as well as the function and specificity of miRNAs (24). Furthermore, editing by ADARs can destabilize dsRNA due to the incorporation of I-U mismatches instead of A-U base-pairs or alternatively, increase its stability by changing A-C mismatches to I-C base-pairs (119). In this way, it is possible that complex structural rearrangements are promoted by editing of specific RNA regions (24,119).

Adenosine deaminases acting on tRNA (ADATs) constitute a family of proteins conserved from yeast to humans that was first identified based on sequence similarity to ADARs (20,33). ADATs catalyze A-to-I editing at or near the anticodon position of tRNAs. Bacterial homologs of the ADAT family, tRNA adenosine deaminases (TadA), also exist (27,33).

Polymerase "Stuttering" Editing

Paramyxoviruses insert guanosine (G) residues into mRNA in order to use alternative reading frames (14,15,120,121). This editing event is a programmed insertion that occurs at a single site in the 15-kilobase viral genome during transcription (10). This editing event arises due to the pseudo-templated "stuttering" of the viral RNA-directed RNA polymerase: the reiterative copying of a template base due to the pausing and backtracking of the viral polymerase and a realignment of the template and

the mRNA (15). The formation of the 3' poly-A tails of viral RNAs also occurs by the pseudo-templated addition of A residues, suggesting that stuttering is an intrinsic enzymatic feature of these polymerases (14,15).

C-Insertion and Dinucleotide Insertion Editing

Virtually all mitochondrial transcripts of the mold *Physarum polycephalum* are edited to some extent by the co-transcriptional insertion of non-encoded nucleotides (29,35,36,122,123). Most editing events involve the site-specific addition of single C residues but U and dinucleotide insertions (UU, AA, GU, GC, and UA) are also observed (10). Although the mechanism by which nucleotides are added at an editing site is unknown, it is thought that the mitochondrial RNA polymerase is responsible for the addition of non-templated nucleotides in response to the presence of *trans*-acting proteins as well as *cis*-acting sequence and/or structural elements in the DNA template, the nascent RNA or both (10).

tRNA Deletion/Insertion Editing

The mitochondrial genome of *Acanthanoeba* encodes 15 tRNAs, 12 of which contain mismatches in the first three base-pairs of the aminoacyl acceptor stem. These mismatches are post-transcriptionally edited to canonical Watson-Crick base-pairs by a process involving the removal of the mismatched nucleotides and the 3'-5' resynthesis using the 3' side of the acceptor stem as a template for nucleotide addition (10,25,26). Because of the polarity of nucleotide addition, it is thought that the polymerizing activity

is not catalyzed by a conventional polymerase but by a nucleotidyltransferase similar to histidine tRNA guanylyltransferase, which adds a G residue to the 5' end of $tRNA_{HIS}$ (26,28,31).

Gene Expression in Trypanosome Mitochondria

kDNA

The mitochondrial genome of trypanosomes, or kDNA, was the first extra-nuclear DNA described (9,124,125). The kDNA is readily visible in trypanosome cells stained with any DNA-specific dye as a disk-shaped spot in the periflagellar region of the mitochondrion. Electron micrographs show it as a giant network of interlocked minicircles and maxicircles, a highly ordered structure that is still poorly understood (Fig. 4) (125).

The mitochondrial genome of *T. brucei* is fragmented in two parts: maxicircles and minicicles. It is composed of a few dozen large (~20 kb) maxicircles and about 5,000 minicircles of ~1 kb in length (2,9,126). Maxicircles are the counterpart to other eukaryotic mitochondrial genomes, encoding components of mitochondrial electron transport complexes and two ribosomal RNA genes. All the coding sequences are concentrated in a contiguous 17 kb stretch called the conservative region (11), while the rest of the molecule is composed of repeated sequences of variable complexity. Minicircles encode guide RNAs that direct editing of mRNAs.

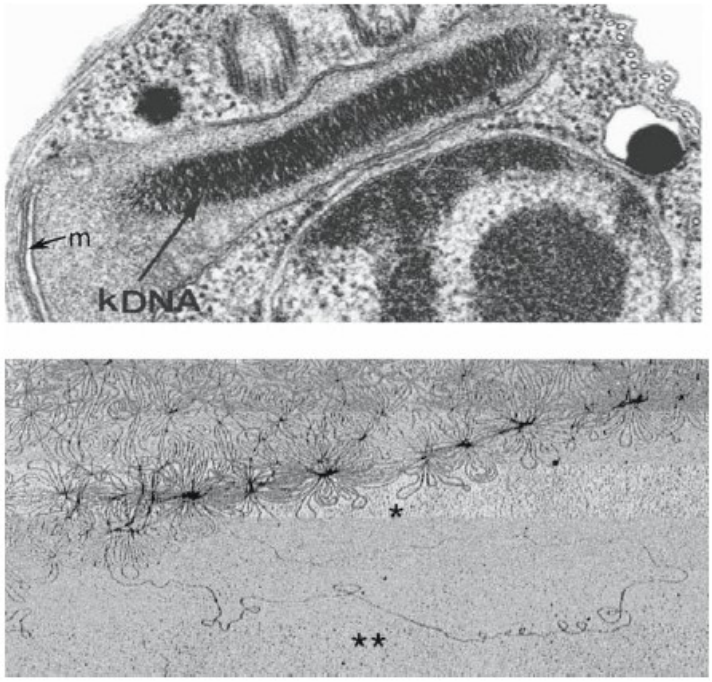

Figure 4. Electron micrographs of kDNA from *T. brucei*. *Top,* The mitochondrial DNA of *T. brucei* is a highly compact structure. *Bottom,* Purified kDNA showing network of catenated minicircles*, and maxicircles**. (Used, with permission, from (9)).

Transcription and Processing of Mitochondrial RNAs

Both maxicircle and minicircle genes are transcribed from unknown promoters by a phage-like, single-subunit mitochondrial RNA polymerase (125,127). Maxicircle genes are transcribed as polycistronic precursors that are then cleaved by a yet-unidentified endoribonuclease into monocistronic rRNAs and pre-mRNAs which are then subject to 3' polyadenylation and editing (9,66,128,129).

The length of the poly-A tail correlates with the editing state of the mRNA and determines its stability (128,130,131). Pre-edited and partially edited mRNAs have short poly-A tails of ~20 nt, whereas fully mature mRNAs have both short poly-A tails and long (~250 nt) poly-A/U tails (11). Polyadenylation destabilizes mRNAs in plant

mitochondria, chloroplasts and bacteria (132,133). In trypanosome mitochondria, however, polyadenylation plays a dual role in RNA stability. Short poly-A tails are essential for the stability of never-edited and edited mRNAs, whereas long poly-A/U tails are added to the short A-tails after completion of the editing process (128).

Trypanosome minicircles contain three to four transcriptional units flanked by 18-bp inverted repeats, which are thought to function in gRNA expression, as transcription initiates 31-32 bp downstream of the 5' repeat at a conserved 5'-RYAYA-3' motif (127,134,135). The function of these 18-bp inverted repeats in gRNA transcriptional initiation, termination or processing is unknown. Since gRNAs contain a 5'-triphosphate end, it was thought that each gRNA was a primary transcript and that their genes contained intrinsic transcriptional initiation and termination sites (135). While this is the case for the relatively rare maxicircle-encoded gRNAs (136), recent evidence suggests that, like other rRNA-and pre-mRNA-encoding maxicircle transcripts, minicircle-encoded gRNA genes are also transcribed polycistronically (66,134) (and B.R. Madina and Cruz-Reyes, unpublished observations). However, it is unknown whether polycistronic gRNAs are processed into a single gRNA or multiple gRNAs. Evidence suggests the existence of a mitochondrial activity that accurately processes the 5'-most gRNA within a polycistron by cleaving its 3' end (134). However, there is no evidence for processing of the 5' end of downstream gRNAs into mature gRNAs.

Thus, the current model of minicircle transcription involves transcriptional initiation at the 5' end of each gRNA transcription unit that results in transcription of a polycistronic precursor containing multiple gRNAs. This polycistronic precursor is

processed at the 3' end of the 5'-most gRNA by a protein complex of approximately 19S to produce a monocistron, while downstream sequences are degraded (134).

Translation in Mitochondria

Although the mitochondrion of trypanosomes may be among the best studied organelles in terms of the replication and maintenance of mitochondrial DNA, transcription, RNA editing and other RNA processing events (11,137), translation in this organelle remains a poorly understood process. For example, while the composition of mitochondrial ribosomes of *T. brucei* and *L. tarentolae* has been described (138-140), no functional studies have been performed. Until recently, several unique features of mitochondrial translation in trypanosomes hindered the study of this process in these organisms.

Cycloheximide inhibits cytosolic translation of most eukaryotes, allowing the study of the prokaryotic-like, chloramphenicol-sensitive translational apparatus in organelles. However, about 10% of cytosolic translation in trypanosomes is resistant to cyclocheximide, which makes the examination of mitochondrial translation very challenging (141,142). The basis of cytosolic cycloheximide-resistant translation is unclear, but it is possible that a fraction of cytosolic ribosomes in trypanosomes either contain subunits that are post-translationally modified, or alternatively are bound by ribosome-associated factors that restrict their binding to cycloheximide. In addition, trypanosome mitochondrial-encoded proteins are extremely hydrophobic and migrate aberrantly under standard electrophoretic conditions; thus far, their separation is only

possible by using a two-dimensional gradient blue-native/SDS-polyacrlyamide gel electrophoretic system (43,142,143).

Ribonuclease III Enzymes

A majority of transcripts are processed via cleavage reactions in order to mature into functional RNAs and, eventually, every RNA is degraded into mononucleotides (133,144). RNA maturation and decay reactions are carried out by ribonucleases, a diverse group of enzymes that cleave RNA through various mechanisms. Ribonucleases can be divided into two broad classes based on the way in which they cleave RNA. Endoribonucleases cleave RNA at internal phosphodiesters while exoribonucleases cleave RNA chains from the 3' or the 5' ends. In recent years, the ribonuclease (RNase) III family of enzymes has gained recognition as a major player in the control of gene expression in a wide variety of organisms (32,145-147). RNase III enzymes play especially prominent roles in trypanosome U-insertion/deletion RNA editing (42).

The RNase III Family

Members of the RNase III family specifically cleave double-stranded RNA (dsRNA) in a magnesium-dependent manner (148-151). All members of this family contain a conserved ribonuclease domain, called the "RNase III" domain. RNase III proteins vary in size, from 200 to >2,000 amino acids, and have been divided into three classes based on their domain composition (Fig. 5) (147,152). Class 1 members contain

a single RNase III domain and a dsRNA-binding domain (dsRBD). Class 2 proteins have two RNase III domains, usually called RNase IIIa and IIIb, as well as a dsRBD. Class 3 proteins, also known as the Dicer family, contain two RNase III domains, a dsRBD and an N-terminal DExH/D-box helicase, a domain of unknown function (DUF283) and a PAZ (Piwi, Aronaute and Zwill) domain (145).

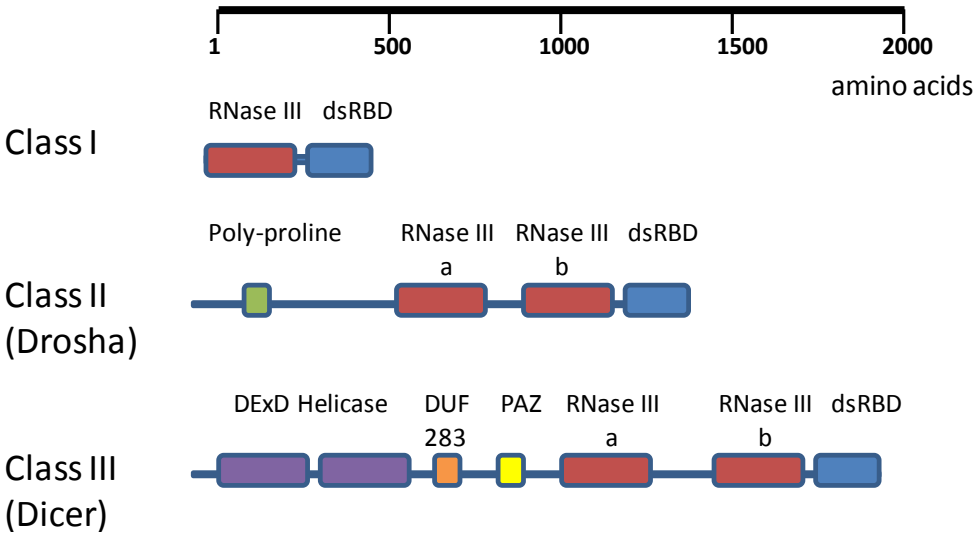

Figure 5. Domain architecture of RNase III classes. Top bar indicates scale in amino acid residues. Domains are indicated. (Adapted from [147]).

Class 1

Class 1 may be the best understood group of the RNase III family and much of the information obtained from studies of members of this class forms the basis of our knowledge of other classes (146). Members of this class are found in bacteria, bacteriophage, fungi and trypanosome mitochondria.

Bacterial RNase III

RNase III from *E. coli* plays important roles in processing rRNA precursors (32,153) and in the regulation of translation by acting on target mRNAs (154-157). Members of class 1 RNase III proteins function as homodimers (158,159), with the RNase III domain mediating dimerization. Biochemical evidence suggests that the two RNase III domains form a single processing center and that each domain is responsible for the hydrolysis of one strand of the duplex substrate (149). In the co-crystal structure of a bacterial RNase III and a dsRNA product, the two active sites form a long "catalytic valley" ((152), Fig X), containing two RNA-binding motifs (RBMs; RBM3 and RBM4) which are conserved in other RNase III family members (Fig. 6) (145,148). Additionally, the dsRBD binds the RNA duplex on the opposite side of the RNase III domain. The dsRBD contains two additional conserved RBMs (called RBM1 and RBM2) that interact with the sugar-phosphate backbone of the bound RNA (160).

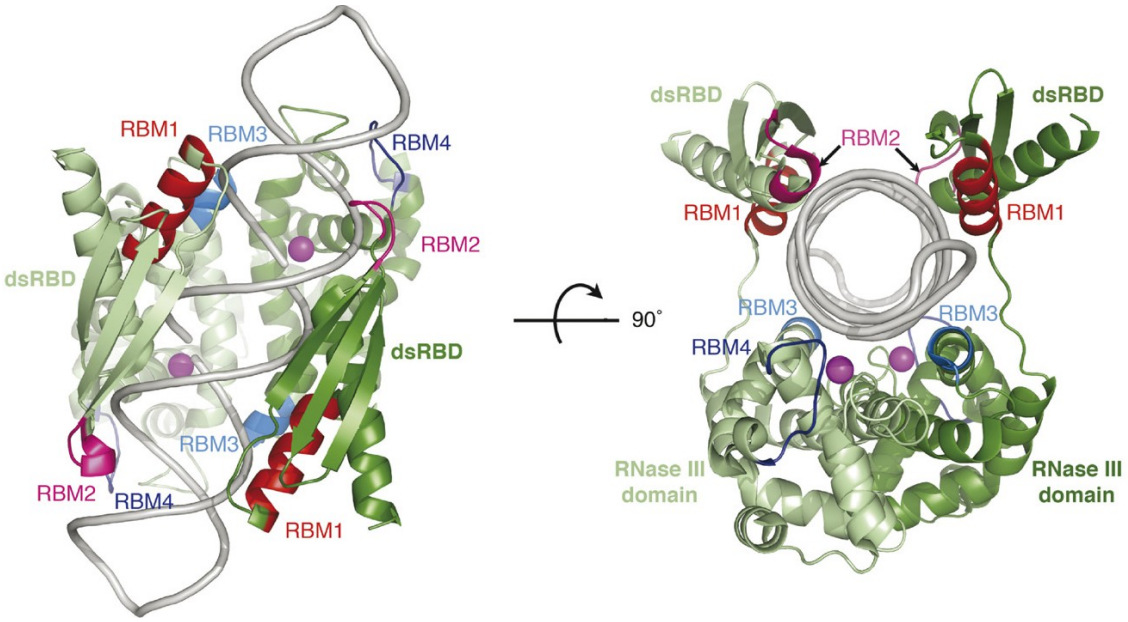

Figure 6. Crystal structure of bacterial RNase III bound to cleaved dsRNA. The two identical polypeptide chains are colored in different shades of green, Mg^{2+} ions are depicted as purple spheres. RBM1 and RBM2 are shown in red; RBM3 and RBM4 are shown in blue. (Adapted, with permission, from (145)).

Rnt1

Rnt1 from *S. cerevisiae* is the best-characterized eukaryotic class 1 RNase III. It plays roles in processing pre-rRNAs, small nuclear RNAs, small nucleolar RNAs and some mRNAs (147,161-165). Substrates for Rnt1 are composed of a dsRNA hairpin ending in a tetraloop (a type of RNA secondary structural motif) containing a consensus AGNN sequence (159,162). The hairpin structure is recognized by the dsRBD through contacts at minor, major, and tetraloop minor grooves on one side of the dsRNA helix which positions the RNase III domains onto the cleavage site, 13-16 base-pairs away from the tetraloop (COSB Fig3).

RNA Editing Endonucleases

The first enzymatic step in RNA editing of mitochondrial transcripts of trypanosomes is catalyzed by the RNA editing endonucleases (RENs) members of the RNase III family (42). Three REN proteins, named REN 1-3, share a similar domain organization but are functionally specialized to cleave specific editing substrates. REN1 cleaves substrates for U-deletion, REN2 cleaves substrates for U-insertion and REN3 exclusively cleaves the cis-guided COII mRNA (57-59).

The first evidence that specialized endoribonuclease activities were present in purified RECC was the differential effect of adenosine nucleotides on the cleavage of deletion and insertion substrates *in vitro* (81). Millimolar concentrations of adenosine diphosphate (ADP) were required for deletion cleavage but partially inhibited insertion cleavage activity. However, it was unclear if this was due to multiple enzymes or one enzyme switching between different states.

The identity of the essential RNase III proteins, REN1, REN2 and REN3, as the deletion, insertion, and cis-guided endonucleases, respectively, was established by combining reverse genetic approaches with *in vitro* assays of RNA cleavage by mitochondrial extracts derived from mutant, knock-out or RNAi-targeted cells (57-59). These proteins contain a N-terminal U1-like Zinc finger, a central RNase III domain that is indispensable for catalysis and a double-stranded RNA binding domain (dsRBD) (46,54,61). These studies, however, did not address the basis for the ability of REN proteins to discriminate deletion, insertion or cis-editing sites, respectively. It is possible that substrate specificity is conferred by protein and substrate determinants and/or by

protein-protein interactions that modulate catalysis or form a favorable architecture. Mutation of conserved residues in the RNase III domain abolished activity, suggesting that these enzymes may employ a similar catalytic mechanism to *E. coli* RNase III (150,166) and, like bacterial counterparts, may homo- and/or heterodimerize (149,167). However, there may be significant deviations from the bacterial prototype enzyme. For example, the 2'-hydroxyl on the ribose immediately upstream of the scissile phosphodiester is absolutely required for REN2 activity, suggesting the involvement of the substrate 2'-hydroxyl in catalysis, in contrast with bacterial RNase III (95,168). In addition, most RNase III enzymes cleave their substrates on both strands, yet REN 1-3 cleave at only one strand in their substrate, which may reflect their heterodimerization with another polypeptide with a non-functional RNase III domain, as shown for the bacterial enzyme (167). For example, two other RECC subunits, KREB4 and KREB5, contain a diverged RNase III domain in which residues required for catalysis are mutated (54,60). It is possible that REN proteins heterodimerize with these inactive RNase III-like proteins and this restricts cleavage to one strand.

In summary, although the editing endoribonucleases are members of the ribonuclease III family and are proposed to follow an RNase IIIl-like cleavage mechanism, significant departures are probable and could potentially be exploited for the development of therapeutic agents as the endoribonucleolytic cleavage of mRNA at an editing site is the first enzymatic step of the editing pathway and must be highly regulated in order to prevent aberrant cleavage of non-target RNAs.

Class 2: Drosha

Drosha is the prototypical member of class II RNase III proteins. It is involved in processing of human pre-rRNA (169) and plays a critical role in the maturation of microRNAs (170) by cleaving pre-miRNAs from primary transcripts, which are then cleaved into mature miRNA duplexes by Dicer (171-173).

Despite the lack of structural information about Drosha, biochemical evidence suggests that it functions as a monomer, with its RNase III domains forming an internal "dimer" structure (174-176). *In vitro*, purified Drosha non-specifically cleaves dsRNA; substrate specificity and correct positioning of the Drosha nuclease center is obtained from a Drosha-interacting protein, DGCR8 (DiGeorge syndrome Critical Region 8) in humans (145,175). DGCR8 is an RNA-binding protein that recognizes dsRNA-ssRNA junctions and positions the Drosha RNase III domains one turn of a dsRNA helix, or 11 base-pairs, away (COSB Fig 4). In this sense, the function of Drosha is analogous to the function of the dsRBD of Rnt1 (and the Dicer PAZ domain, see below). The Drosha-DGCR8 complex is named the "microprocessor" to indicate its role in processing of pri-miRNAs (175). DGCR8 contains a proline-binding WW domain that is thought to interact with the proline-rich N terminus of Drosha, raising the possibility that other WW domain-containing proteins may associate with Drosha and confer alternative specificity.

Class 3: Dicer

Class 3, or Dicer, RNase III enzymes cleave dsRNA into products of 21-27 nucleotides in length (177) which participate in RNA silencing pathways. Human Dicer has a complex domain architecture; it is composed of a DExH/D-box helicase domain, a domain of unknown function (DUF283), a dsRBD, a PAZ domain and two tandem RNase III domains (145). Dicer is often thought to be a "molecular ruler" that recognizes the ends of dsRNA substrates and cleaves them a specific distance away, generating a dsRNA product of discrete size. The PAZ domain binds one end of dsRNA (COSB 53,54-56) (178-181) and a "connector helix" sets the measuring distance from the end of the dsRNA molecule to the cleavage site (182). Thus, like Rnt1 and Drosha-DGCR8, accessory domains in RNase III proteins target the RNase III domain to the correct site in a dsRNA molecule. The functions of the other domains in Dicer are still unknown.

DExH/D-box Proteins

Helicases may be the most prevalent class of enzymes in nucleic acid metabolism (183,184). They function *in vivo* at virtually every step in gene expression including DNA replication, repair, recombination, transcription, ribosome biogenesis and RNA processing, translation and decay (185,186). Helicases couple the free energy released from the hydrolysis of nucleoside triphosphates to their translocation along nucleic acids, the unwinding/separation of helical segments of nucleic acids or, in some cases,

disruption of protein-nucleic acid interactions (113,183,184,186-188). Because of their myriad and often essential functions, helicases are ubiquitous and evolutionarily conserved proteins.

Classification of Helicases

Helicases are characterized and classified by the presence of short conserved amino acid sequence motifs (184,185,189). They are grouped into three large (SF1-3) and two small (SF4 and 5) superfamilies based on the number of motifs, their amino acid sequence and spacing (Fig. 7A). SF1 and SF2 have at least seven conserved motifs (I, Ia, II, III, IV, V and VI), have a catalytic core that consists of two domains separated by a linker region and function as monomers (Fig. 7B), whereas SF3-SF5 assemble into hexamers, with a core of six individual domains arranged in a ring. These domains are similar to the ATP-binding core of the recombination protein RecA. The conserved helicase motifs involved in NTP binding and hydrolysis are similar to the Walker A and B boxes of ATP Synthase. Other conserved motifs are involved in coupling the NTP hydrolysis state to conformational changes in the protein and nucleic acid binding. The core structure of helicases permits cycles of nucleic acid binding and release driven by NTP binding and hydrolysis. Current models describe the function of helicases as the coupling of NTP hydrolysis to the directional translocation/unwinding of nucleic acid helical structures (190).

Helicases are generally classified as DNA or RNA helicases according to their nucleic acid polymer specificity, although some helicases can unwind both and others

unwind RNA-DNA duplexes (184). DNA helicases are involved in DNA replication, repair and recombination. RNA helicases are involved in all facets of RNA metabolism.

Most RNA helicases belong to the DExH/D-box group of SF2 proteins (191,192). DExH/D-box proteins derive their name from the single-letter code of the four amino acids of motif II (corresponding to the Walker B motif of ATPases). The DExH/D-box group of proteins is further classified into several families including DEAD-box, DEAH-box, and DExH-box families which are distinguished by sequence differences beyond motif II. DEAD-box members constitute the largest subgroup and demonstrate significant functional differences compared to other helicases (185,191). The *in vitro* activities of DExH/D-box proteins suggest that they function as major players in remodeling ribonucleoprotein complexes: RNA unwinding activity has been demonstrated for several DExH/D proteins, while some can function as RNPases (disrupting protein-RNA interactions) and others have RNA annealing activity (113).

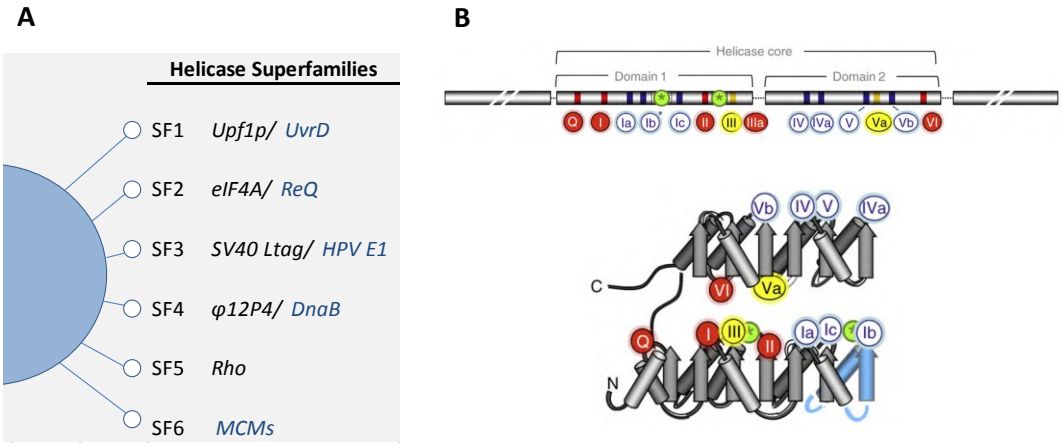

Figure 7. Helicase superfamilies and conserved helicase motifs. A) Classification of helicase superfamilies according to (184); representative superfamily members are indicated, RNA helicases are indicated in black, DNA helicases in blue. B) *TOP:* Characteristic sequence motifs of SF2 helicase core. Coloring indicates function, red: ATP binding and hydrolysis; blue: nucleic acid binding; yellow: coordination of nucleic acid and ATP binding;. B) *BOTTOM:* Location of helicase motifs in helicase core fold (arrows represent β-sheets; cylinders represent α-helices) (Adapted, with permission, from (186)).

DExH/D-box Proteins as Regulators of RNA Metabolism

The unwinding activity of the catalytic core of helicases is not sequence-specific (186,190). In contrast to the lack of specificity shown *in vitro*, most RNA helicases function in specific processes in the cell and many seem to be specialized for a particular process (183). Thus, *in vivo*, helicases employ mechanisms to recognize and to assemble onto the correct substrate. In addition, their activities must be tightly regulated to prevent potentially deleterious consequences.

Perhaps the best understood *in vivo* functions of eukaryotic RNA helicases are those of Prp22p, and eIF4A-III (193,194). Prp22p allows the release of spliced mRNA from the spliceosome and also participates in the exon ligation step (194). Prp22p changes its position on the RNA several times in an ATP-dependent manner that likely involves its directional movement in a 3'-5' direction along the mature mRNA which dissociates several RNA-RNA and RNA-protein contacts, causing the release of the mRNA from the spliceosome (Fig. 8A) (186).

Although eIF4A-III unwinds dsRNA *in vitro*, in the cell it functions as an adaptor of the exon junction complex, a multi-component complex that binds upstream of exon-exon junctions and remains stably associated to the mRNA after export into the cytoplasm, controlling its localization, translation and decay (Fig. 8B) (193,195). This stable binding is achieved by arresting the eIF4A-III ATP hydrolysis cycle by preventing the dissociation of hydrolysis products by EJC components Magoh and Y14 (Fig. 8C) (196).

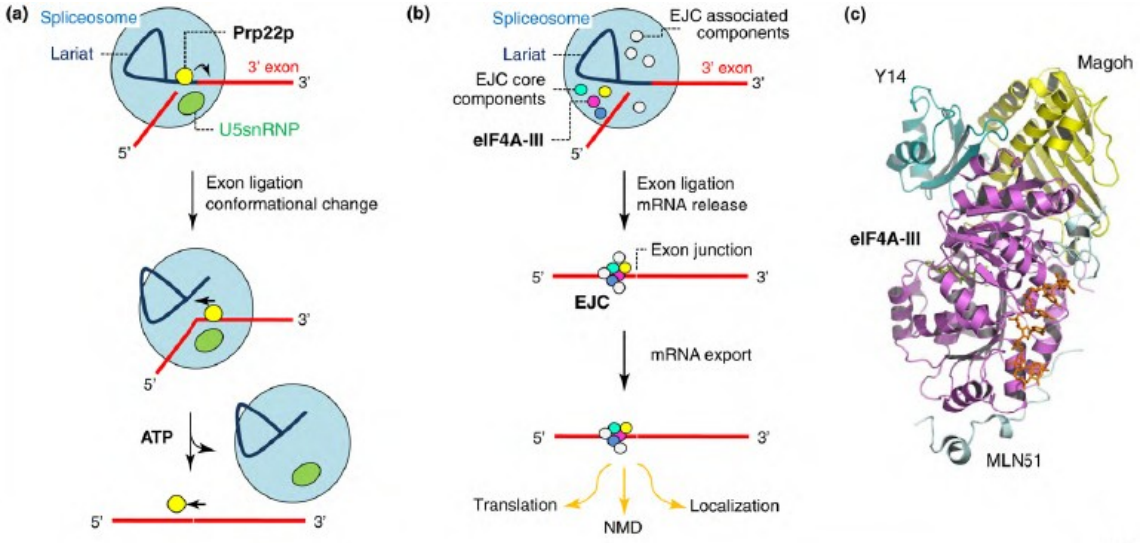

Figure 8. Functions of Prp22p and eIF4A-III. (A) Prp22p the releases of mRNA from the spliceosome (194). Prp22p binds close to the 3' splice site. After exon ligation, Prp22p changes its position and binds 3' to the exon junction, moves towards the 5' end in an ATP-dependent manner, removing the spliceosome. (B) eIF-4A-III acts as an adaptor for other components of the EJC. Core components remain stably bound to eIF4A-III, while other components associate and dissociate at various stages in mRNA metabolism. (C) Structure of the core EJC bound to RNA (197). Proteins are indicated and RNA is shown in orange. Obtained, with permission, from (186).

Regulation of RNA Helicase Activities in Protein Complexes

RNA helicases usually function in the context of multi-protein complexes where their enzymatic activities may be regulated by binding of co-factors (185). Stimulation of RNA helicase activities often reflects the ability of the binding partner to increase the affinity of the helicase for RNA, which might be a general strategy for the recruitment of helicases to a particular substrate.

The activities of Dbp5p, a DEAD-box protein that function in mRNA export, are spatially activated by its association with Gle1p, a protein that accumulates on the

cytoplasmic side of the nuclear pore and the small molecule inositolhexakisphosphate (198,199)

Binding partners can also inhibit helicase activities (for example, eIF4A-III in the EJC), RNA binding or their association to other proteins. Binding of the nucleoporin NUP214 blocks the RNA binding site of Dbp5, thus inactivating it (200). The tumor suppressor protein, PDCD4 inhibits translation by binding to the translation initiation factor eIF4A, preventing its association with initiation factor eIF4G (199,201,202).

Dissertation Overview

This dissertation is composed of two main parts. In the first part (Chapters II-IV), studies are presented that are aimed at understanding the mechanism of substrate recognition by RECC in trypanosome RNA editing. We show the identification of RNA-protein interactions between purified RECC and RNA substrates *in vitro* and the determinants for their association. In the second part (Chapters V and VI), I present evidence for the involvement of two novel mitochondrial DExH-box proteins in RNA editing.

Chapter II shows specific RNA-protein interactions identified by cross-linking between purified RECC and the first deletion editing site of A6 pre-mRNA. We present evidence for the selectivity of these interactions using competition analysis with homologous and heterologous RNAs. In addition, the identified cross-linking interactions also show specificity for unpaired residues 5' of the pre-mRNA/gRNA

anchor duplex, a hallmark of editing sites. We propose that the RNA-protein interactions we observed may be involved in the recognition of editing sites and/or catalysis.

Chapter III describes determinants for full-round editing and RECC interaction at editing site 2 of A6 pre-mRNA, an insertion site. We show that RECC recognition of substrate duplexes flanking the editing site is sequence-independent and that these duplexes can be minimally composed of a single helical turn. Substitution of ribose 2' functionalties defined 2'-hydroxyls in the gRNA loop and editing site-flanking helices that are required for the cleavage of pre-mRNA and association with RECC. Notably, a single 2'-hydroxyl immediately upstream of the editing site is indispensable for cleavage but not for cross-linking with RECC.

Chapter IV defines relatively simple single- and doble-stranded substrate determinants for the association and gRNA-directed RNA cleavage through the use of a direct binding assay in parallel with cross-linking and enzymatic experiments. Cleavage of RNA requires a duplex with gRNA of approximately 15 base-pairs in addition to a 5' overhang of 12 nt. A second element in either the substrate or guide strand is required unless longer duplexes are used. Importantly, we show that the cleavage and substrate binding activities of purified RECC can be uncoupled and that a single-stranded RNA requirement for association can be upstream or downstream of the duplex. In addition, we show that the endonuclease REN1 cross-links to editing sites and we propose a model whereby one or more RNase III-type endonucleases present in RECC mediate the initial binding and scrutiny of potential ligands.

Chapter V describes the characterization of a novel DExH-box RNA helicase, named RNA editing-associated helicase (REH) 2. We show that REH2 forms unique ribonucleoprotein complexes that bind gRNA and exhibit double-stranded RNA-unwinding activity. The REH2 helicase motif I and a conserved double-stranded RNA-binding domain are required for the integrity of REH2 RNPs and their helicase and gRNA-binding activities. REH2 associates, via RNA, with RECC, along with a wide array of RNA editing accessory factors and mitochondrial ribosomes. A model is proposed in which REH2 RNPs are integral components of RNA-linked supramolecular networks that orchestrate the expression of the mitochondrial genome of trypanosomes.

Finally, Chapter VI presents the conclusions derived from the studies above and some possible future directions.

CHAPTER II

RNA EDITING COMPLEX INTERACTIONS WITH A SITE FOR FULL-ROUND U DELETION IN *TRYPANOSOMA BRUCEI**

Summary

Trypanosome U-insertion and U-deletion RNA editing of mitochondrial pre-mRNAs is catalyzed by multisubunit editing complexes as directed by partially complementary guide RNAs. The basic enzymatic activities and protein composition of these high-molecular mass complexes have been under intense study, but their specific protein interactions with functional pre-mRNA/gRNA substrates remains unknown.

We show that editing complexes purified through extensive ion-exchange chromatography and immunoprecipitation make specific cross-linking interactions with A6 pre-mRNA containing a single ^{32}P and photoreactive 4-thioU at the scissile bond of a functional site for full-round U deletion. At least four direct protein–RNA contacts are detected at this site by cross-linking. All four interactions are stimulated by unpaired residues just 5' of the pre-mRNA/gRNA anchor duplex, but strongly inhibited by pairing of the editing site region.

Furthermore, competition analysis with homologous and heterologous transcripts

* Reprinted with permission from "RNA editing complex interactions with a site for full-round U deletion in *Trypanosoma brucei*" by A. Sacharidou, C. Cifuentes-Rojas, K. Halbig, A. Hernandez, L.J. Dangott, M. de Nova-Ocampo and J. Cruz-Reyes. 2006. *RNA*. 12:1219-1228.

suggests preferential contacts of the editing complex with the mRNA/gRNA duplex substrate. This apparent structural selectivity suggests that the RNA–protein interactions we observe may be involved in recognition of editing sites and/or catalysis in assembled complexes.

Author Contributions

Pages 53-79 are the original text and figures reported in reference #94. RNA editing complex interactions with a site for full-round U deletion in Trypanosoma brucei. Sacharidou A, Cifuentes-Rojas C, Halbig K, Hernandez A, Dangott LJ, De Nova-Ocampo M, Cruz-Reyes J. RNA. 2006 Jul;12(7):1219-28. Author contributions are as follows:

Pg. 62, Fig. 9: Sacharidou A

De Nova-Ocampo M

Pg. 64, Fig. 10: Sacharidou A, Halbig K

Pg. 66, Fig. 11: Sacharidou A

Pg. 68, Fig. 12: Sacharidou A

Pg. 70, Fig. 13: Sacharidou A

Pg. 72, Fig. 14: Hernandez A

Pg. 73, Fig. 15: Sacharidou A.

Dangott LJ helped with some protein work in the analyses.

Introduction

Mitochondrial mRNAs in trypanosomatid protozoa including *Trypanosoma, Leishmania*, and *Crithidia* species undergo a unique form of RNA editing by cycles of uridylate insertion or deletion at numerous editing sites (ESs). This post-transcriptional mRNA maturation progresses with a general 3'–5' polarity and is catalyzed by a large multisubunit RNA editing complex (also termed 20S editosome or L-complex) proposed to contain between 8 and 20 polypeptides depending on the purification protocol (46,47,71,203,204). The smaller number presumably reflects high-stringency purification conditions and tight association of the subunits in the resulting complexes.

Partially complementary guide RNA (gRNA) transcripts direct this process, which is believed to initiate with the formation of an ''anchor duplex'' with pre-mRNA. Catalysis of a single editing cycle involves three basic activities, namely, mRNA endonuclease, 3' terminal uridylyltransferase (Tutase in insertion) or 3' to 5' U-specific exoribonuclease (in deletion), and RNA ligase. So far, the known catalytic subunits in the editing complex include a TUTase (KRET2, also termed LC-6b; (47,65), a U-specific exonuclease (KREP6, LC-2; (63), two RNA ligases (KREL1, Band IV, LC-7a and KREL2, band V, LC-9; (68,205,206), deletion and insertion endonucleases (KREN1 and KREN2; (58,59), respectively), and an endonuclease/exonuclease (KREPA3, band VI, LC-7b; (207)). All these protein subunits have been cloned and characterized *in vitro* and *in vivo*. A significant amount of information has been obtained on the structural and functional composition of editing complexes (for reviews, see (54,55,208); however, the specific RNA–protein interactions in assembled complexes during

recognition of pre-mRNA/gRNA duplex substrates and catalysis of full editing cycles are unknown.

Several reported protein subunits contain conserved motifs for nucleic acid binding, but only a purified recombinant KREPA3 has been shown to exhibit RNA-binding activity (207). In addition to core essential subunits, a few auxiliary components involved in editing are known, including the annealing factors MRP1 (gBP21) and MRP2 (gBP25) (100,209-211), and the gRNA-binding factor RBP16 (103). Other proposed factors are an RNA helicase, REAP1, and TbRGG1 (46,107,114,212). All factors mentioned above are either weakly or not associated with editing complexes and dispensable for *in vitro* editing (46,47,203,213).

Here, using photocross-linking we report four protein interactions in intimate contact with the first editing site (ES1) for full-round U deletion in an A6 pre-mRNA/gRNA substrate that co-purify and co-immunoprecipitate with editing complexes. All four RNA–protein cross-links exhibit structural selectivity for the single-stranded character of the editing site region. Together, the data indicate that the cross-linking events described here are mediated by one or more stably bound core subunits. To our knowledge, this is the first report of specific RNA–protein interactions of editing complexes with a functional site for full-round RNA editing.

Methods

Pre-mRNA and gRNA Substrates

The ATPase 6 (A6) pre-mRNA editing substrates (Seiwert et al.1996) for deletion with gRNA D33 (51) were prepared as previously described. The site-specific radio-labeled and 4-thioU modified pre-mRNAs were obtained by ligation of two fragments as in (214). For bond 45 (ES1), the acceptor and donor RNAs were 5'-GGAAAGGUUAGGGGGAGGAG-AGAAGAAAGGGAAAGUUGUGAUUU-3' and 5'-**U**GGAGUUAUAGAAUACUU-ACCUGGCAUC-3', the latter containing a 5'-terminal 4-thioU in bold. For bond 34 (ES4), 5'-GGAAAGGUUAGGGGGAGGAGAGAAGAAAGGGAAAG-3' and 5'-UUGUGAUUUUGGAGUUAUAGAAUACUUACCUGGCAUC-3'; and bond 51, 5'-GGAAAGGUUAGGGGGAGGAGAGAAGAAAGGGAAAGUUGUGAUUUUG-GAGU-3' and 5'-UGGAGUUAUAGAAUACUUACCUGGCAUC-3' were used, respectively. The acceptor RNAs were transcribed using the Uhlenbeck single-stranded T7 transcription method (215) and gel-purified. The donor thiolated RNAs were chemically synthesized by Dharmacon. The 4-thioU residue of the donor piece was radiolabeled to high-specific activity with polynucleotide kinase and [γ-^{32}P]ATP (using a 1:2 molar ratio of 5' ends:ATP), gel-purified, and ligated to the acceptor piece using the following DNA oligonucleotide bridges (bond 45): 5'-TATTCTATAACTCCAAAATCACAACTT-TCC-3'; (bond 34), 5'-

AACTCCAAAATCACAACTTTCCCTTTGTTC-3'; (bond 51), 5'-GCCAGGTAAGTATTCTATAACTCCAAAATC-3'. A 3:1:2 molar ratio of acceptor/donor/bridge molecules was used.

Preparation of Crude Mitochondrial Extracts and Fractions Containing Enriched or Purified Editing Complexes

Procyclic form (Pf) *T. brucei* strain TREU667 was grown in Cunningham media, and mitochondrial crude extracts were prepared as in (216) with modifications as in (89). Mitochondrial crude extracts were fractionated by ion-exchange chromatography in consecutive Q-sepharose (Q1) DNA-cellulose (D), and Q-sepharose (Q2) columns, as described by (203) and (89). The elution fractions with the peak of editing complexes determined by Western blot analysis or editing activity also contained the peak of cross-linking activity in all purification steps.

Editing, Adenylylation, and Cross-linking Analysis

Full-round editing reactions assembled in 20-μL mixtures with pre-annealed 3'-end labeled A6 pre-mRNA (~10 fmol) and gRNA D33 (~1.2 pmol) and adenylylation of RNA ligases in editing complexes were performed as in (81,82) and (217), respectively. For photocross-linking analysis, editing reactions were assembled as above, but in the absence of nucleotides, which somewhat improves cross-linking. The complete mixtures were incubated for 10 min at 26°C and an additional 10 min. on ice prior to irradiation with 365-nm UV light (on ice for 10 min, ~5 cm below a Spectroline 150-V

lamp) and subsequent treatment with Rnases A and T1 (50 mg/mL and 120 U/μL) for 10 min at 37°C. After addition of 7 μL of 4X Laemmli buffer, the samples were analyzed by SDS-PAGE and autoradiography. RNA competitors at the indicated molar excess were included in the reaction mixture supplemented to the pre-annealed pre-mRNA/gRNA duplex in both cross-linking and editing assays. The 15-nt homopolymers were synthesized by IDT. The 121-nt viral RNA H121 was a gift from Cheng C. Kao (218). We have determined in native gels that our pre-annealing step yields >95% of the pre-mRNA in a duplex with gRNA D33 (not shown), so further gRNA addition in Fig. 6A,B should hybridize virtually all pre-mRNA.

Immunoprecipitation and Western Blot Analysis

Immunoprecipitations were performed essentially as described by (204) with minor modifications. For immunoprecipitation analysis of cross-linking proteins, editing reactions were scaled up 10 times and cross-linked as described above. One hundred micro-liters of Immunomagnetic beads (Dynabeads M-450; Dynal) were coupled with 225 μL of monoclonal antibodies (kindly provided by the laboratory of Ken Stuart, SBRI Seattle) and 1% BSA. Editing reactions were incubated with antibody-coated beads for 1 h at 4°C using a bi-directional shaker and occasional tapping. After washing two times with 100 μL of immunoprecipitation buffer (10 mM Tris at pH 7.2, 10 mM $MgCl_2$, 200 mM KCl, 0.1% Triton-X 100) the beads were resuspended with 100 μL of TE buffer and incubated in the presence of Rnases A and T1 as described above. Upon the addition of 30 μL of 4X Laemmli buffer, the bead

suspension was boiled at 100°C for 5 min and the supernatant analyzed by SDS-PAGE and autoradiography. The entire 200 μL unbound fraction and 100 μL washes mixed with 60 μL and 30 μL of 4X Laemmli buffer, respectively, boiled as well as analyzed. For Western blot analysis with the indicated monoclonal antibodies, protein samples (cross-linked to RNA or not) were separated by SDS-PAGE, blotted, and probed with the indicated mouse monoclonal antibodies at a dilution of 1:25–1:50. The secondary antibody was applied at a 1:5,000 dilution and the blot developed using the ECL plus system (Amersham).

Results

To search for RNA–protein interactions in assembled RNA editing complexes, we generated a 72-nt A6 pre-mRNA substrate containing a single ^{32}P and 4-thioU at the scissile bond of the first editing site (ES1) for U deletion (Fig. 9A). Prior to photo-crosslinking, this thiolated pre-mRNA was pre-annealed with gRNA and mixed with editing complex preparations, as in standard *in vitro* reactions (see Materials and Methods). Importantly, the thiolated pre-mRNA supports accurate *in vitro* deletion of three uridylates as directed by the partially complementary gRNA D33 (51), although slightly less efficiently than unmodified pre-mRNA (Fig. 9A, B). This indicates that the presence of a thio-uridylate immediately 3' of the scissile bond does not significantly interfere with editing activity.

We initially utilized crude mitochondrial lysate that was fractionated by Q-sepharose chromatography (Q1 column; Fig. 9C) to detect protein cross-links to ES1. This column has been previously used to enrich active editing complexes (203,204,219). Several cross-links of various intensities are evident across the fractionated lysate upon irradiation with 365-nm UV light. At least four of them, at about 40, 50, 60, and 100 kDa, appeared to closely co-purify with editing complexes as detected by immunoblots of known core subunits, particularly in the peak fractions 9–11 (‘‘peak Q1 fractions’’; Fig. 9C,D).

Other prominent cross-links were detected at about 75, 150, and 250 kDa in or near these fractions. Proteinase K inactivation of all cross-links in the peak Q1 fractions showed that they are protein dependent (not shown), so from here onward we will refer to them as p40, p50, p60, and p100. The peak Q1 fractions eluted away, between 150 and 200 mM KCl, from most proteins in the mitochondrial crude extract, and therefore appear significantly enriched (Fig. 9E). These fractions were pooled and further purified by two subsequent steps of ion-exchange chromatography in DNA-cellulose and Q-sepharose columns, respectively (Fig. 10; data not shown).

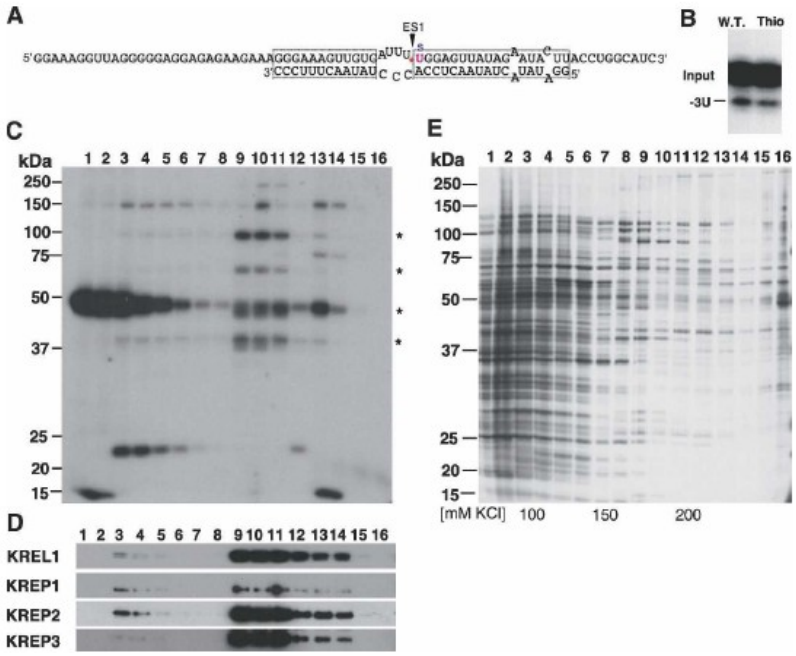

Figure 9. RNA–protein interactions detected by photocross-linking co-purify with RECC in Q-sepharose fractionated mitochondrial extract. (A) Diagram of the 72-nt A6 pre-mRNA substrate annealed with 33-nt gRNA D33. Boxes indicate predicted duplexes flanking ES1. The 4-thioU (sU) and 5'-^{32}P-radiolabed bond (*) are positioned at the double-strand/single-strand junction that defines ES1 (arrowhead). (B) U-deletion *in vitro* assay of unmodified (W.T.) and thiolated A6 pre-mRNA paired with gRNA D33. Input and accurate -3U deletion RNAs are indicated. (C) UV irradiation (365 nm) of the pre-mRNA/gRNA substrate with Q-sepharose fractions. The asterisks indicate the positions of four proteinase K-sensitive cross-links that co-purify with RECCs. RECC was detected in immunoblots (D) of four known subunits KREL1 (also termed TbMP52), KREP1 (MP8), KREPA2 (MP63), and KREPA3 (TbMP42). RECC and co-purifying crosslinks peak in fractions 9–11. The molecular size is indicated in kilodaltons (E) Silver staining of Q1-sepharose fractions. The peak Q1 fractions elute between 150 and 200 mM KCl.

Notably, p40, p50, p60, and p100 co-purify with editing activity in both columns. However, additional bands are detected in the DNA-cellulose column (‘‘D’’) peak fractions, although not reproducibly in our protein preparations (not shown). The peak fractions of the second Q-sepharose column (‘‘Q2’’ fractions 13–15) show primarily p40, p50, p60, and p100 (Fig. 10A), precisely co-purifying with isolated silver-stained polypeptides and full-round deletion activity (Fig. 10B, C). Notably, our peak Q2 fractions exhibit a pattern of major stained protein bands, plus a few additional fainter bands (Fig. 10B), that is remarkably similar to that of editing complexes purified with either the same protocol (89) or another biochemical purification strategy (204). Both the same protein pattern and relative intensity of individual bands are conserved whether silver or SYPRO Ruby staining is used (data not shown). Importantly, p40, p50, p60, and p100 co-localize with stained bands in the Q2 peak fractions (Fig. 10D). Furthermore, these cross-links are only detected if the targeted residue is thiolated and, therefore, upon 365-nm but not 260-nm UV light irradiation (data not shown).

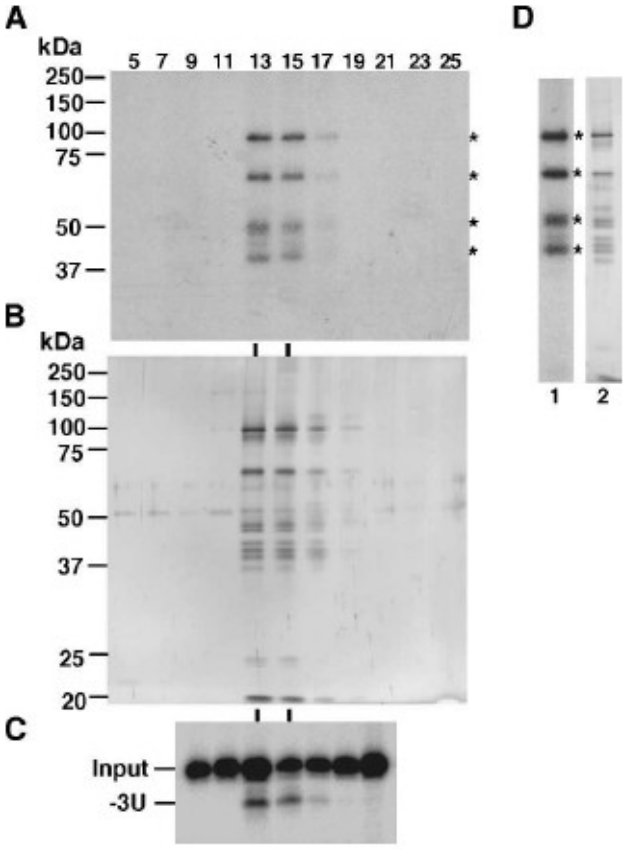

Figure 10. p40, p50, p60, and p100 co-purify with RECC after extensive ion-exchange chromatography. Peak Q1 fractions from Figure 1 were fractionated on DNA-cellulose (D) and a second Q-sepharose column (Q2). Shown are the relevant odd fractions of the Q2 elution. The four protein–RNA cross-links (A) precisely co-purify with silver-stained RECCs (B) and U-deletion activity (C). The U-deletion activity of Q2 fractions was assayed at the ES1 of the 3'-end-labeled A6 pre-mRNA. The RNA input and accurate deletion product (-3U) are indicated. (D) The four protein–RNA cross-links (lane 1) co-localize with silver-stained protein components (lane 2) of the peak Q2 fractions (no. 13–15).

A direct comparison of the protein content and cross-linking pattern of Q1, D, and Q2 peak fractions (Fig. 11A,B) indicates that the four RNA–protein interactions described above are conserved throughout the purification of active editing complexes (Fig. 11C) and most likely involve the same proteins. Other cross-links previously observed in Q1 and occasionally in D fractions are significantly reduced or lost in Q2 fractions. The peak Q2 fractions (13–15) contain ~1/6,000 of the original crude mitochondrial extract protein and exhibit a simpler protein pattern than the parental D and Q1 fractions. This extent of purification is consistent with others reported using similar protocols (203,204,220). There is at least a ~10-fold further purification compared to the whole-cell protein content; however, the specific activity of editing complexes could not be calculated since the *in vitro* editing assay is not linear with protein added, particularly in cruder fractions (203,204,220); data not shown). Together, these data suggest that p40, p50, p60, and p100 are tightly associated with purified active editing complexes and that they make intimate contacts with the targeted editing site.

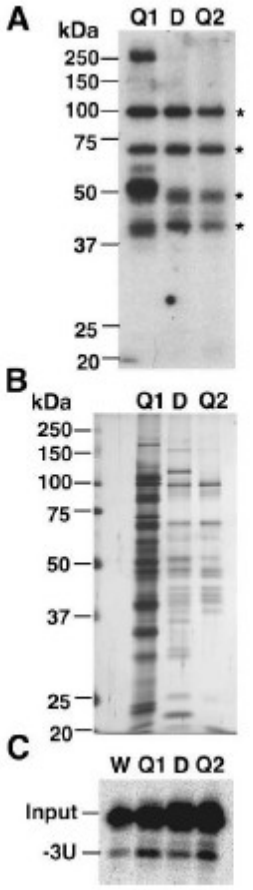

Figure 11. Side-by-side gel analyses of Q1, D, and Q2 peak fractions. (A) protein–RNA cross-linking interactions, (B) silver staining, and (C) full-round U-deletion activity. The latter includes a lane with the original whole mitochondrial extract (W).

To further confirm this association, we performed co-immunoprecipitation assays (co-IP) using monoclonal antibodies that are known to immunoprecipitate active editing complexes (204,221).

Analysis of the peak Q1 fraction shows efficient co-IP of the p40, p50, p60, and p100 kDa cross-links by anti-KREPA3 antibodies (Fig. 12). Relative to a control lane showing the starting cross-linked sample (‘‘C’’), the unbound lane (‘‘U’’) shows a significant decrease in three cross-links, p40, p60, and p100, and their corresponding enrichment in the bound material (‘‘B’’) after two washes (‘‘W2’’). Most cross-linking activity at ~50 kDa remains in the unbound fraction, but a significant amount (above background levels) co-Ips with the editing complex, as compared with a mock assay with no antibodies. We interpret this as indicative of at least two proteins comigrating at ~50 kDa, one corresponding to a stably bound component (p50) of editing complexes and another representing a mitochondrial protein that is presumably abundant but not tightly associated with editing complexes. Consistent with this notion, the latter cross-link may account for the prominent ~50-kDa band in the flow through and first few fractions in the initial chromatographic step (Fig. 9C), and apparent trailing into the peak editing fractions. The same cross-linking protein is significantly reduced or lost in the D and Q2 peak fractions (Fig. 11A), and in most gels, it appears to migrate slightly above the proposed p50 cross-link (e.g., Figs. 1C, 3A, 4).

Co-IP assays were also performed with antibodies against two other editing subunits, KREPA2 and KREL1, and in cases, p40, p50, p60, and p100 selectively immunoprecipitate with editing complexes (not shown). Additional analyses were

performed to confirm the specificity of the p40–100 interaction with editing complexes. These include a positive control showing efficient co-IP of radiolabeled RNA ligase subunits (via ^{32}P-adenylylation; (204,221)) and a negative control with a nonrelated antibody (not shown). The virtual absence of the ~150- and ~250-kDa cross-links in Q2 fractions (Fig. 10A) and their reduction to near background levels in co-IP assays (Fig. 12) suggest that the cross-linking proteins are either weakly or not bound to editing complexes.

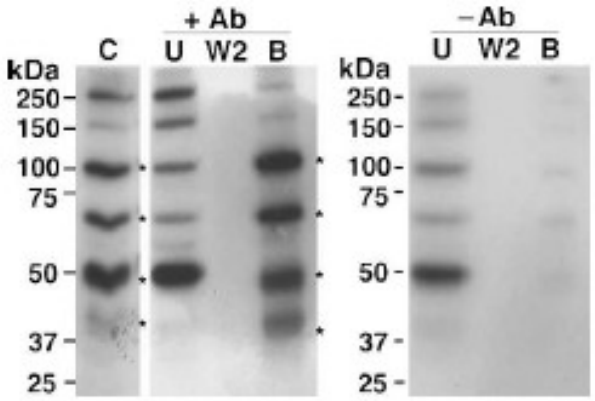

Figure 12. p40, p50, p60, and p100 co-immunoprecipitate with RECCs. Protein–RNA cross-links in a peak Q1 fraction before (C lane) and after a co-IP assay with anti-MP42 antibodies (+Ab), including the unbound (U), second wash (W2), and bound immunoprecipitated (B) fractions. A parallel mock co-IP assay with no antibodies (-Ab) is shown.

Combined, our extensive chromatography purification and immunoprecipitation analyses show at least four RNA–proteins cross-links between one or more stably bound subunits of editing complexes and a site for full-round deletion in an A6 substrate. Notably, these cross-links specifically target the [^{32}P]-labeled photoreactive 4-thioU positioned at the scissile bond of this functional substrate.

To determine whether or not the polypeptides that bind ES1 also contact other positions of the A6 pre-mRNA/gRNA substrate, we moved the [^{32}P]-labeled photo-reactive 4-thioU a few nucleotides away from the scissile bond at ES1 (bond 45; Fig. 13A). In one case, we tested the upstream bond 34 that corresponds to the second deletion site (ES4) in the natural A6 substrate, and in another, the downstream bond 51 in the never-edited region of this transcript. Both positions are located within the predicted upstream and downstream duplexes formed by the partially complementary gRNA D33, respectively (Fig. 13A, top and middle RNA pairs). Notably, all four protein–RNA interactions detected by cross-linking at functional ES1 (bond 45) are absent at either duplex position (Fig. 13B). This suggests that the observed RNA–protein cross-linking interactions may exhibit structural selectivity for single-strandedness of the editing site. To confirm this apparent preference for single-stranded residues adjoining the photo-reactive 4-thioU, we annealed the pre-mRNA to a gRNA derivative (31.dx) that extends the upstream and downstream duplexes into a single contiguous duplex (Fig. 113A, bottom pair). We found that base-pairing of the ES1 region with 31.dx strongly inhibits all cross-links observed with the parental gRNA D33 (Fig. 13C).

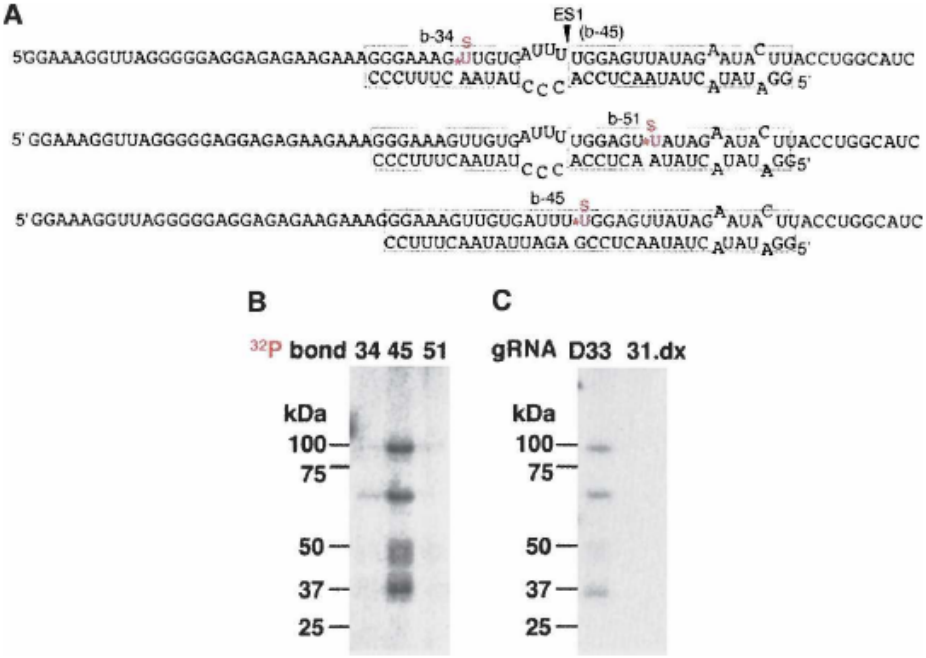

Figure 13. All four RNA–protein interactions detected by cross-linking in Pf editing complexes are favored by single-strandedness at the editing site. (A) Diagrams of A6 pre-mRNA/ D33 pairs as in Fig. 1A, but with the [5'- 32P] thiolated U at upstream (b-34) or downstream (b-51) bonds (top and middle RNA pairs, respectively). The position of ES1 (b-45) is also indicated. The A6 pre-mRNA modified at b-45 was also paired to a gRNA D33 derivative (31-dx) that forms a continuous duplex across ES1 (bottom pair). (B) Parallel crosslinking assays in a Q2 peak fraction of radiolabeled pre-mRNA at each of three indicated bonds above, paired with gRNA D33. (C) Cross-links of pre-mRNA modified at b-45 and annealed with either D33 or a D33-like derivative (31-dx) that fully base-pairs the ES1 and directs no deletion.

Together, our data indicate that all four cross-linking proteins observed at ES1 are favored by the single-strand character of the editing site. Importantly, precise gRNA base-pairing across ES1 inhibits *in vitro* U deletion at this site (48).

To assess the specificity of the interaction between editing complexes and A6 pre-mRNA/D33 substrate, we supplemented the cross-linking assay with a molar excess of various non-radiolabeled RNA competitors (Fig. 14A–C). Interestingly, addition of 10- and 25-fold excess (relative to radiolabeled A6 pre-mRNA) of the homologous A6 pre-mRNA virtually abolished all cross-linking (Fig. 14A, lanes 1–3), whereas another pre-mRNA (Cyb; lanes 4–5) and tRNA (lanes 6–7) were only slightly inhibitory at the same concentration. The partial effect of the latter heterologous competitors seems specific to these transcripts, as further addition (25-fold) of gRNA D33 did not affect the crosslinking efficiency (lanes 8–9). Note that the assay includes pre-mRNA ((51); see Methods section). The inhibition by the A6 pre-mRNA competitor is consistent with its ability to base-pair with gRNA D33. Additional heterologous transcripts including the noncomplementary gRNA gRPS12, viral RNA H121 (25- to 50-fold excess), and several homopolymers (100-fold excess) were slightly or not inhibitory (Fig. 14B, C; data not shown). Up to 100-fold further addition of gRNA D33 (i.e., ~200-fold excess overall) in the latter assays was not inhibitory (Fig. 14B, lanes 5, 6).

We also tested the above RNA competitors on full-round U deletion. As expected, the homologous pre-mRNA was fully inhibitory at 25-fold excess, whereas all other competitors in Fig. 14A–C were little or not inhibitory at the same concentration (Fig. 6D; data not shown). Combined, the above competition analyses on cross-linking

and editing assays suggest that editing complexes may be able to distinguish the pre-mRNA/gRNA duplex from individual substrate strands and from nonrelated structured or relatively nonstructured transcripts. Additional studies are currently under way in our laboratory to further address this question.

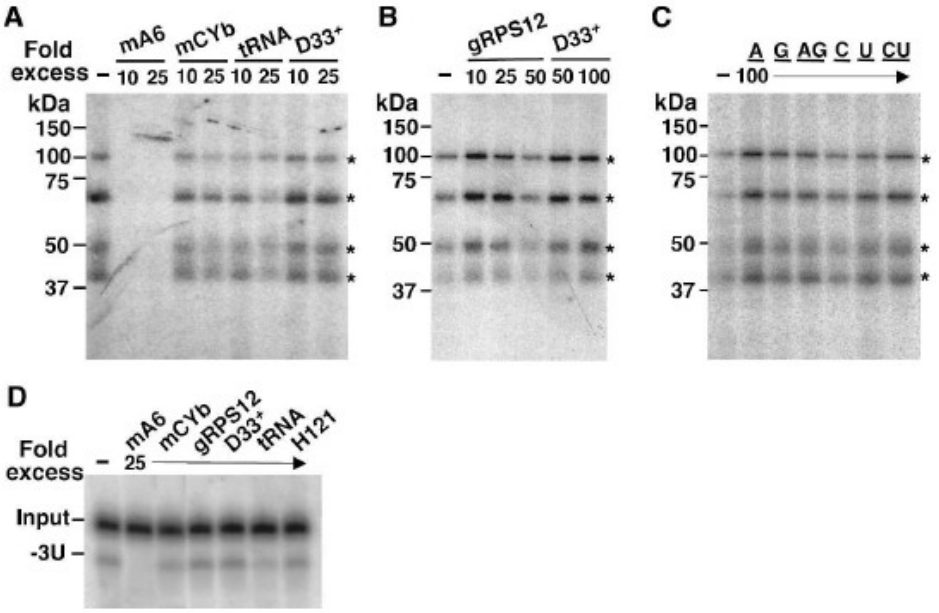

Figure 14. Homologous and heterologous RNA competitors in cross-linking and editing assays. Cross-linking with or without (A) 10 and 25-molar excess of homologous A6 pre-mRNA (mA6) or heterologous Cyb pre-mRNA (mCYb) and tRNA, or complementary gRNA D33. (+) Additional D33 over the standard amount (~100-fold excess) present in the cross-linking assay. (B) Ten-, 25-, and 50-fold excess of noncomplementary gRNA gRPS12 or 50- and 100-fold excess of complementary gRNA D33 (over its standard level in the assay, as in A). (C) Hundred-fold excess of 15-nt oligomers. (D) U-deletion assay with or without 25-fold excess of the indicated transcripts (~125-fold overall in the case of gRNA D33).

Based on the observed gel mobility of p40, p50, p60, and p100, we suspected that one or more of them could correspond to known subunits of editing complexes. To test this possibility, we transferred the reactions to a membrane after cross-linking and performed Western analysis using available monoclonal antibodies to identify the co-localizing proteins. Our initial analysis showed a precise co-localization between p60 and KREPA2 (~60 kDa; band III; LC-4), whereas p40 did not precisely match with KREPA3 (~40 kDa; band VI; LC-7b) (Fig. 15). Furthermore, p40 and p50 do not comigrate with the editing RNA ligases (^{32}P-labeled by adenylylation; (217); data not shown). MS analyses of the protein bands matching the crosslinks are under way, but due to the possibility of cross-contamination between similar-size subunits (particularly in the ~90–100 kDa and ~40–55 kDa size ranges; (54)) additional work using epitope-tagging of candidate subunits will be required to establish definite subunit assignments for p40, p50, and p100, and confirm that p60 corresponds to KREPA2.

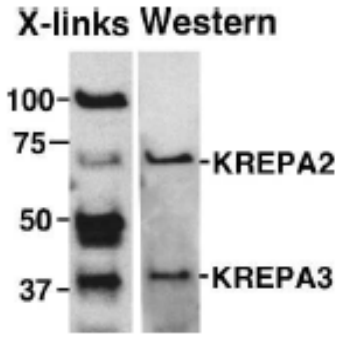

Figure 15. p60 colocalizes with the KREPA2 subunit. A cross-linking reaction (X-links lane) and subsequent Western blot analysis of the same gel (Western lane) with anti-KREPA2 and anti-KREPA3 antibodies.

Overall, the extensive biochemical co-purification and co-immunoprecipitation of p40, p50, p60, and p100 with active editing complexes indicates that the cross-links involve one or more stably bound components of editing complexes. Moreover, our analysis of substrate features and response to RNA competitors suggests that editing complexes and possibly these particular RNA–protein interactions exhibit structural selectivity for the editing substrate used in our studies.

Discussion

The specific RNA–protein interactions in editing complexes that lead to their activation and catalysis of faithful RNA editing cycles in trypanosomes are unknown. The purpose of this study was to identify specific pre-mRNA/protein contacts using assembled editing complexes and an A6 pre-mRNA/gRNA substrate for full-round editing *in vitro*.

We found at least four protein interactions, p40, p50, p60, and p100, in direct contact with ES1 for U deletion. These interactions revealed by protein–RNA cross-linking involve one or more tightly bound subunits of editing complexes since they precisely co-purify with editing activity upon extensive ion-exchange chromatography in three consecutive columns and co-IP using monoclonal antibodies raised against known editing complex subunits. The ion-exchange chromatography (89) and immunoprecipitation (68,204,221) approaches applied in this study were previously exploited to efficiently purify active editing complexes and study their protein

composition. All major protein components of the complexes originally observed by (203) are also present in the complexes prepared by immunoprecipitation and similar chromatography or affinity purifications (46,47,204,221).

The identification of the cross-linking polypeptides reported is evidently necessary to begin dissecting their potential in editing. The protein banding pattern of our purified editing complexes is remarkably similar to others previously reported using related biochemical purification schemes (46,203,204), and associations between specific subunits and protein bands in those patterns have been proposed (for reviews, see (54,55)). Based on the co-localization of p60 with band III (KREPA2; LC-4) in both silver-stained gels (Fig. 10D) and immunoblots (Fig. 15; data not shown) we speculate that p60 may indeed correspond to band III. The precise molecular function of this subunit has not been defined, but it has been found associated with KREPC2 and KREL1 in a purified subcomplex that catalyzes partial (precleaved) deletion editing (67). These authors have speculated that KREPA2 could use its potentially regulatory OB-fold to coordinate the sequential enzymatic steps of U deletion. Furthermore, this subunit has also been proposed to play a critical structural role in the formation or stability of entire editing complexes (77,222). Other reported subunits of predicted molecular size similar to p60, although not found during the peptide sequencing of band III (by Edman degradation; (222)), include KREN2 and KREPB2, an essential insertion-specific endonuclease and a potential endonuclease, respectively (58,59). At least the essential KREN2 is expected in our purified complexes, either migrating with band III

(possibly at sub-stoichiometric levels) or near to it. Another reported subunit, KRET2, appeared to be sub-stoichiometric (71) in similarly purified complexes.

P100 precisely colocalizes with the prominent band I (203), which corresponds to an (~99 kDa) exonuclease proposed to function in U deletion (KREPC2; LC-3; (54,55). However, we cannot exclude the possibility that p100 may be the closely migrating KREN1, an essential U deletion-specific endonuclease (46) expected in our purified active complexes, or alternatively KREPC1 (~100 kDa), a candidate editing exonuclease (46) potentially present in our preparation. Any of the above likely p100 candidates is consistent with our search for subunits that bind and cross-link a deletion site.

Several known editing complex subunits could account for the p40 and p50 cross-links we observe (54,55), including five (~41- to 49-kDa) subunits with a conserved U1-like zn-finger domain potentially involved in macromolecular interactions with RNA substrates or other proteins in the complex. Two of these proteins also exhibit a C-terminal Pumilio RNA-binding domain and less conserved RNase III motifs potentially involved in endonuclease cleavage. Our Western blot analysis revealed that p40 is not KREPA3 (~42 kDa; Fig. 15). Moreover, the RNA ligases KREL1 (~52 kDa) and KREL2 (~45 kDa) migrate between the p40 and p50 cross-links in high-resolution acrylamide gels and therefore are different proteins (not shown). It is also conceivable that one or more of these proteins, p40, p50, and/or p100, correspond to novel subunits of editing complexes. Further work is under way to identify these proteins and their potential roles in deletion.

KREPA3 (~42-kDa subunit) and five related subunits exhibit apparent zn-finger domains and/or an OB fold. The former are found in many regulatory proteins and could mediate interactions with nucleic acids or with other proteins, whereas the latter typically provides a nonspecific binding platform for single- and double-stranded nucleic acids (76). KREPA3 is the only subunit known so far to bind RNA (207). Surprisingly, a recombinant version of this protein was reported to exhibit endonuclease and (3'–5') exonuclease activities on a stretch of unpaired uridylates in a partial RNA hybrid, although KREPA3 lacks recognizable nuclease domains. While these activities are editing-like, the substrate used in that study is not functional, and the proposed protein–RNA interaction remains to be confirmed in assembled editing complexes. RNAi knockdown of KREPA3 does not appreciably disassemble editing complexes, but reduces *in vivo* and *in vitro* editing (207). Thus, the reported properties of rKREPA3 suggest that this subunit has important roles in editing. Whether or not KREPA3 is functionally similar or even redundant to any structurally related subunit remains to be determined. KREPA3 was not detected in our analysis at ES1, however this may reflect a limitation of our ''zero-distance'' cross-linking approach. That is, even if a protein specifically binds the targeted site, the thiolated uridylate and adjacent amino acid side chain may not be properly orientated with each other for efficient photoreaction.

A double-strand/single-strand junction just 5' of the downstream ''anchor'' duplex is a critical feature of functional editing sites (48,50). Interestingly, the cross-links we observe are strongly inhibited by gRNA base-pairing of the editing site (Fig. 13). This observation suggests that the p40–100 interactions with the substrate exhibit

structural selectivity for the mismatched preedited ES1, but are inhibited by gRNA complementarity across the edited site. In addition to simple mRNA/gRNA mismatches at editing sites, structural studies have indicated that other features of functional pre-mRNA/gRNA pairs may determine the basis for endonuclease recognition (223,224). Nevertheless, it is feasible that p40, p60, p50, and p100 may play important roles during recognition and/or catalysis at editing sites. A previous study of U insertion in Leishmania proposed that two RNA cross-linking proteins, ~80 and 100 kDa, from highly enriched editing extracts may be associated with editing site recognition, but the RNA substrate positions cross-linked remain to be determined (225).

Our competition analyses also suggest that editing complexes may preferentially recognize features of the pre-mRNA/gRNA hybrid (Fig. 14). gRNA D33 is supplemented at ~100-fold the level of the radiolabeled A6 premRNA, in both standard cross-linking and editing assays, although we have seen that a ~200-fold excess affects neither activity (Fig. 14B,D). Importantly, we have seen in native gels that during the pre-incubation step in our assays virtually all radiolabeled A6 pre-mRNsA anneals to gRNA D33 (see Materials and Methods section; data not shown). Addition of non-radiolabeled A6 pre-mRNA at 10-fold excess (or less) strongly inhibits cross-linking and editing (Fig. 14A; data not shown), whereas 25- to 100-fold excess of other transcripts that should not hybridize with gRNA D33 have little or no effect. Interestingly, significantly structured transcripts such as tRNA (25-fold) appear relatively more inhibitory than predicted low-structured sequences, including the gRNA constructs (50-fold) and short RNA homopolymers (100-fold) tested (Fig. 14; data not shown). This

apparent binding preference of editing complexes for RNA substrates *in vitro* is under further investigation in our laboratory.

Our observation of multiple cross-linking interactions at the ES1 for deletion in the A6 pre-mRNA/gRNA substrate may reflect that this site is dense with protein contacts in editing complexes (possibly not all detected by our cross-linking approach). Also the natural dynamics of interacting subunits, variable RNA substrate conformations, or protein breakdown may account for the multiple cross-links detected. These possibilities will be further studied in our laboratory. Furthermore, we observed the same cross-linking pattern in immunoprecipitated editing complexes enriched from bloodstream form trypanosomes ((115); data not shown). Together with our extensive purification of the procyclic complexes, this suggests that these proteins are part of the core complex and may not directly account for developmental regulation.

Finally, editing complexes contain subgroups of apparently related subunits sharing similar conserved motif (54). This may reflect the proposed functional and structural partition of insertion and deletion components in editing complexes (67,73,74,81,82), and functions outside editing, including polycistronic mRNA, gRNA, and rRNA processing (129,134). Whether the editing complex cross-links reported here and/or other subunits occur at different deletion or insertion sites and in other substrates is currently under investigation in our laboratory.

CHAPTER III

SUBSTRATE DETERMINANTS FOR RNA EDITING AND EDITING COMPLEX INTERACTIONS AT A SITE FOR FULL-ROUND U INSERTION*

Summary

Multisubunit RNA editing complexes catalyze uridylate insertion/deletion RNA editing directed by complementary guide RNAs (gRNAs). Editing in trypanosome mitochondria is transcript-specific and developmentally controlled, but the molecular mechanisms of substrate specificity remain unknown.

Here we used a minimal A6 pre-mRNA/gRNA substrate to define functional determinants for full-round insertion and editing complex interactions at the editing site 2 (ES2). Editing begins with pre-mRNA cleavage within an internal loop flanked by upstream and downstream duplexes with gRNA. We found that substrate recognition around the internal loop is sequence-independent and that completely artificial duplexes spanning a single helical turn are functional.

Furthermore, after our report of cross-linking interactions at the deletion ES1 (35), we show for the first time editing complex contacts at an insertion ES. Our studies

* Reprinted with permission from "Substrate determinants for RNA editing and editing complex interactions at a site for full-round U insertion" by C. Cifuentes-Rojas, P. Pavia, A. Hernandez, D. Osterwisch, C. Puerta and J. Cruz-Reyes. 2007. *The Journal of Biological Chemistry*. 202 (7):4265-4276.

using site-specific ribose 2' substitutions defined 2'-hydroxyls within the (a) gRNA loop region and (b) flanking helices that markedly stimulate both pre-mRNA cleavage and editing complex interactions at ES2. Modification of the downstream helix affected scissile bond specificity. Notably, a single 2'-hydroxyl at ES2 is essential for cleavage but dispensable for editing complex cross-linking. This study provides new insights on substrate recognition during full-round editing, including the relevance of secondary structure and the first functional association of specific (pre-mRNA and gRNA) riboses with both endonuclease cleavage and cross-linking activities of editing complexes at an ES. Importantly, most observed cross-linking interactions are both conserved and relatively stable at ES2 and ES1 in hybrid substrates. However, they were also detected as transient low-stability contacts in a non-edited transcript.

Author Contributions

Pages 80-117 are the original text and figures reported in reference #95. Substrate determinants for RNA editing and editing complex interactions at a site for full-round U insertion. Cifuentes-Rojas C, Pavia P, Hernandez A, Osterwisch D, Puerta C, Cruz-Reyes J. J Biol Chem. 2007, 282(7):4265-76. Author contributions are as follows:

Pg. 94, Fig. 16: Cifuentes-Rojas, C.

Pg. 96, Fig. 17:Cifuentes-Rojas, C.

Pg. 100, Fig. 18: Cifuentes-Rojas C, Pavia P.

Pg. 102, Fig. 19: Cifuentes-Rojas C.

Pg. 104, Fig. 20: Cifuentes-Rojas C, Pavia P.

Pg. 105, Fig. 21: Cruz-Reyes J.

Pg. 107, Fig. 22: Cifuentes-Rojas C.

Pg. 109, Fig. 23: Hernandez A.

Pg. 110, Fig. 24: Hernandez A.

Introduction

The single-mitochondrion containing kinetoplastid protozoa, including species of Trypanosoma and Leishmania, use cycles of uridylate insertion or deletion at numerous editing sites (ESs) within pre-mRNAs to generate mature mRNAs (for recent reviews, see (54,55,208)). This post-transcriptional mRNA maturation is catalyzed by a multisubunit editing complex (58,59,63,152,226) with specificity for the ESs being directed by small transacting guide RNAs (gRNAs) that are partially complementary to pre-mRNA (49,226-228).

A significant body of information has been accumulated on the functional and structural composition of editing complexes, including the identity of the subunits catalyzing the three steps of each editing cycle; they are mRNA cleavage at deletion and insertion ESs (58,59), U addition or U removal (63,65,87) and RNA ligation at deletion and insertion ESs (67,68,73,74,205,206). The complexes are heterogeneous in protein composition but share most of the approximately 20 subunits identified (61). Several

factors are also known or proposed to play auxiliary roles in editing (46,100,103,104,107,114,209-212), although they are dispensable *in vitro* (46,47,203,213). Much less is known about the mechanisms of substrate recognition including the protein subunits and substrate determinants that distinguish pre-edited (pre-) mRNAs from other transcripts and DNA in mitochondria. We recently reported the first observations of direct editing complex interactions with a functional site for full-round U deletion, showed preferential association with the editing substrate, and provided evidence for one of the interacting subunits corresponding to KREPA2 (94). However, editing complex interactions at insertion sites have not been reported. Other recent reports showed that bacterially expressed recombinant versions of the subunits KREPA3 and KREPA4 bind RNA (207,229). The latter exhibited specificity for a gRNA 3'-oligo(U) tail.

In pre-mRNA/gRNA substrates, unpaired pre-mRNA uridylates or unpaired gRNA purines are landmarks of deletion or insertion sites, respectively (227) and the number of such residues dictates the extent of U removal or addition (49,50,227). The two kinds of editing are likely to be differentially regulated as they involve separate activities and enzymes (58,59,73,74,81,82) and there is evidence for their physical separation in heterogeneous complexes and subcomplexes (61,67). Interestingly, efficient deletion and insertion editing have distinct requirements for a proposed pre-mRNA/gRNA ligation bridge (51,93) and artificially interconverted sites use differing pre-mRNA lengths (91). The above observations suggest that the editing complex recognitions in and near an ES may also differ between the two editing types. Our

interconversion of functional ESs from deletion to insertion and vice versa experimentally demonstrated that the basic determinants that commit editing complexes into full-cycle deletion or insertion reside within the internal loop containing the targeted ES (91). However, additional features proximal and/or distal to an ES may modulate the efficiency of editosome assembly and catalysis. For example, discrete sequence changes affecting the pairing potential of residues adjoining an ES can significantly impact the specificity and efficacy of full-round and partial ("pre-cleaved") editing (93,230). The current model of trypanosome RNA editing postulates that natural sites should be flanked by a proximal upstream duplex between a purine-rich pre-mRNA sequence and a gRNA 3' poly-U tail (49,226-228) and an adjacent pre-mRNA/gRNA downstream "anchor" duplex that directs cleavage (48-50,227). Mutational analysis of the gRNA 3' region that stabilizes the upstream duplex can significantly enhance full-round editing *in vitro* (51,83). In *Leishmania tarentolae*, an upstream duplex was used in model U-insertion substrates by one group (231,232) but was not essential according to another (219). The latter group proposed that pre-mRNA purine sequences have a role in editing that is independent of base-pairing with gRNA (231,232). In a Cyb pre-mRNA substrate, a 34-nt A/U element appeared to modulate gRNA-directed and gRNA-independent insertion (233) and a discrete 5' determinant near an editing site in a ND7 substrate was proposed (219,232). In *Trypanosoma brucei*, the three model systems that currently recreate a full-round editing *in vitro*, A6, Cyb, and RPS12 (49,50,84,91,93) are based on natural purine-rich pre-mRNA fragments. We showed in A6 constructs that natural pre-mRNA extensions protruding from the upstream and downstream duplexes

could be replaced by unnatural stretches without significant effects on editing. In one such construct, about half of a 5' poly-purine run proposed to stimulate editing *in vitro* (49) was substituted by pyrimidines (91). However, whether or not a specific pre-mRNA (or gRNA) sequence or its natural base composition modulates editing remains unclear. Previous structural studies in solutions of different natural-like mRNA/gRNA pairs have proposed that a common secondary structure may be important for editing complex recognition (234) but this has not been tested in functional *in vitro* systems.

Here we performed systematic sequence mutagenesis and ribose 2'-deoxynucleoside substitutions of a minimal A6 pre-mRNA/gRNA substrate to define functional determinants for both full-round U insertion and editing complex interactions at the targeted ES2. Our competition analyses of editing and RNA-protein interactions showed evidence of preferential association of editing complexes with the hybrid substrate. We observed that the requirement for the duplexes flanking the internal loop is sequence-independent, and artificial helices spanning a single turn support efficient editing. We also found that specific ribose 2 –hydroxyls in both strands of the downstream helix and, surprisingly, in the gRNAloop region strongly stimulate both pre-mRNA cleavage and editing complex interactions at the targeted insertion site. Moreover, 2'-deoxy substitution of the downstream helix affected scissile bond selectivity, whereas the tested modifications in either pre-mRNA or gRNA strand had relatively moderate effects. Notably, the 2 –hydroxyl moiety adjoining the scissile bond is an essential determinant of insertion, potentially involved in cleavage catalysis.

The current studies of trypanosome full-round insertion editing provide significant insights on the relevance of the substrate secondary structure rather than its specific sequence and suggest that specific pre-mRNA and gRNA riboses significantly affect both pre-mRNA cleavage and editing complex interactions at the targeted bond.

Methods

Pre-mRNA, gRNA, RNA/DNA Chimeric, and Photoreactive Substrates

The starting substrate in these studies was the minimized ATPase 6 (A6) 45-nt pre-mRNA (91) paired with a variant of the enhanced gRNA gA6[14]USD-3A (83). This substrate directs full-round insertion of 3Us at ES2 and uses pre-edited ES1 to increase the stability of the downstream duplex. RNAs were transcribed from a DNA template as described by Milligan et al. (215) gel-purified, and quantified using an ND-1000 spectrophotometer (NanoDrop®). The DNA templates below are 3'-extended with the T7 promoter complementary strand TATAGTGAGTCGTATTA. The number of the RNA pair using the transcript product is in brackets (see Fig. 1; #, operational number).

Pre-mRNAs-[2]-CTTTCCCTTTCTTCTCTCCTCCCCCTCCTTTCCCTATAACT-CCAAAATCAGTACATACGCATACATC, #309; [3] CTTTCCCTTTCTTCTCTC-CTCCCCCTCCTTTCCCTATAACTCCAAAATCAGTACATACGCGCCC, #352; [4] CTTTCCCTTTCTTCTCTCCTCCCCCTCCTTTCCCTATAACTCCAA-AATCAGTACATCGCGCCC, #356; [5] CTCCCCCTCCTTTCCCTATAACTCC-

AAAATCAGTACATCGCGCCC,#445; [6] CTATAACTCCAAAATCAGTACATCG-CGCCCTTCCTCCTCCTTTCC,#447; [7] CTATAACTCCAAAATCAGTACATCG-CGCCCTTAAAGAAAGAGCCC,#465; [8] CTTGACTCCAAAATCAGTACATCG-CGCCCTTCCTCCTCCTTTCC,#463; [9] CTGACTCCAAAATCAGTACATCGC-GCCCTTCCTCCTCCTTTCC,#464; [10] CCACACTCACATCAGTACATCGC-GCCCTTCCTCCTCCTTTCCC, #556; [11] CCACATCACATCAGTACATCGCG-CCCTTCCTCCTCCCTTTCCC,#561; [12] GGACATCACATCAGTACATCGCG-CCCTTCCTCCTCCCTTTCCC,#563.

gRNAs-[2] AATGTATGCGTATACTTCGTTTATCTCGGAGTTATAGTATATCC, #307; [3] GGGCGCGTATACTTCGTTTATCTCGGAGTTATAGTATATCC, #349; [4] GGGCGCGATACTTCGTTTATCTCGGAGTTATAGTATATCC, #350; [8] GGGCGCGATACTTCGTTTATCTCGGAGTCTAGTATATCC, #467; [9] GGGC-GCGATACTTCGTTTATCTCGGAGTCAGTATATCC, #468; [10] GGGCGCGAT-ACTTCGTTTATGTGAGTGTGGTATATCC, #557; [11] GGGCGCGATACTTCG-TTTATGTGATGTGGTATATCC, #568; [12 12] GGGCGCGATACTTCGTTTATG-TGATGTCC, #569. Deoxynucleoside-substituted transcripts were made by (IDT, Inc.), and 2'-F and 2'-OCH_3 modified transcripts were by (Dharmacon, Boulder, CO). Ribonucleotides are denoted by the prefix "r".

*Pre-mRNA Strand-[*13] GGGGGAGGAGArGrArArGrArArArGrGrGrArArA-rGrUrArCrUrGrArUrUrUrUrGrGrArGrUrUrArUrArG, #403; [14] rGrGrGrGrGrA-rGrGrArGrAGAAGAAAGGGArArArGrUrArCrUrGrArUrUrUrUrGrGrArGrUrUrArUr ArG, #404; [15] rGrGrGrGrGrArGrGrArGrArGrArArGrArArArGrGrGrArArArGrU-

rArCrUGrArUrUrUrUrGrGrArGrUrUrArUrArG, #405; [16] rGrGrGrGrGrArGrGrA-rGrArGrArArGrArArArGrGrGrArArArGrUrArCrUmGrArUrUrUrUrGrGrArGrUrUrArUrArG, #456; [17] rGrGrGrGrGrArGrGrArGrArGrArArGrArArArGrGrGrArArArG-rUrArCrUfGrArUrUrUrUrGrGrArGrUrUrArUrArG, #519; [18] rGrGrGrGrGrA-rGrGrArGrArGrArArGrArArArGrGrGrArArArGrUrArCrUrGrArUrUrUrUGGAGTTATAGrA, #429; [19] rGrGrGrGrGrArGrGrArGrArGrArArGrArArArGrGrGrArArArG-rUrArCrUrGrATTTTGGAGTTATAGrA, #441; [20] rGrGrGrGrGrArGrGrArGrArG-rArArGrArArArGrGrGrArArArGrUrArCrUrGATTTTGGAGTTATAGrA, #451; [21] GrGrGrGrGrArGrGrArGrArGrArArGrArArArGrGrGrArArArGrUrArCrUrGArUrUrUrUrGrGrArGrUrUrArUrArG, #430; [26] rGrGrGrGrGrArGrGrArGrArGrArArGrArA-rArGrGrGrArArArGrUACTrGrArUrUrUrUrGrGrArGrUrUrArUrArG, #565.

gRNA Strand—[22, 23] rGrGrArUrArUrArCrUrArUrArArCrUrCrCrGrArGrArUrA-rArArC-rGrArA-rGrUrUrUTCCCTTTCTTrU, #485; [24, 25] rGrGrArUrArUrA-CTATAACTCCrGrArGrArUrArArArCrGrArArGrUrUrUrUrCrCrCrUrUrUrCrUrUrU, #487; [26, 28] rGrGrArUrArUrArCrUrArUrArArCrUrCrCrGrArGrArUAAACGA-ArGrUrUrUrUrCrCrCrUrUrUrCrUrUrU, #486.

Photoreactive Substrates—Each pre-mRNA was obtained by ligation of two pieces (235). All thiolated RNA pairs numbers are indicated by a colon. Acceptor pieces: [15’] rGrGrGrGrGrArGrGrArGrArGrArArGrArArArGrrGrArArArGrUrA-rCrUG, #424; [23’] 5’rGrGrGrGrGrArGrGrArGrAGAAGAAAGGGArArArGrUrA-rCrUrG, #524’; [27’,28’] rGrGrGrGrGrArGrGrArGrArGrArArGrArArArGrGrGrArA-rArGrUACTrG, #560; the common donor piece (4-ThioU)rArUrUrUrUrGrGrA-rGrUrUrArUrArG,

#401. For Pair-25' the donor piece was (4-ThioU)- rArUrUrUrUGGAGTTATAGArA, #567). The acceptor pieces were synthesized by IDT®, and the thiolated donors were synthesized by Dharmacon®. The donor pieces were radiolabeled to high specific activity with T4 polynucleotide kinase and [γ-32P]ATP (MP Biomedicals) using a 1:2 molar ratio of ends:ATP, gel-purified, and ligated to the acceptor piece as described (94) using as the bridge CTATAACTCCAAAATACAGTACTTTCCCTTTC, #553. The molar ratio of acceptor/donor/bridge was 2:1:1.5.

Purification of Editing Complexes

Procyclic *T. brucei* strain TREU667 was grown in Cunningham media and mitochondrial extracts were prepared as described (216). Editing complexes were enriched by Q-Sepharose ion exchange chromatography and further purified by DNA-cellulose affinity chromatography as reported (94,203). Additional enrichment can be achieved by using another step of Q-Sepharose; however, both editing and cross-linking activities were equivalent in the minimal substrate for full-round insertion (91) by complexes from the two-step and three-step purifications (see the figure on p. 107 and data not shown). Fractions with the peak of editing activity were used for all the experiments.

Editing and Cleavage Assays and Quantitation Analysis

Full-round U insertion was performed as described (82). Briefly, a 2-μl mixture with pre-annealed 3'-end-labeled pre-mRNA (~10 fmol) and gRNA (1.2 pmol) was

completed to 20 µl with 10 mM MRB buffer (10 mM magnesium acetate, 10 mM KCl, 1 mM EDTA, pH 8, 25 mM Tris-HCl, pH 8, and 5% glycerol), 150 µM UTP, 3 µM ATP, and 2 µl of peak editing fraction. The reaction was incubated at 26 °C for 60 min, deproteinized, and resolved in 9% acrylamide, 7 M urea gels. Editing complexes were pretreated with 10 mM Ppi to score total cleavage in absence of RNA ligase activity (73). Neither ATP nor UTP were added to this assay, and the cleavage products were resolved in 15% PAGE with 7 M urea. Ribonuclease T1 and hydroxyl ladders were used to confirm the cleavage at ES2 (not shown). All pre-mRNAs for editing were 3'-end-radiolabeled with [^{32}P]cytidine 3',5'-(bis)phosphate except for the 2'-F-modified transcript (Pair-17), which had to be made with a 3'-terminal deoxynucleoside (Dharmacon). Such a terminus prevents radiolabeling with T4 RNA ligase (236) so this transcript was 5'-end-labeled with T4 polynucleotide kinase. Data were visualized by phosphorimaging and/or x-ray autoradiography, and quantitation was performed using a STORM PhosphorImager (ImageQuant 5.0, GE Healthcare). Each panel in the figures corresponds to one of two replica series performed simultaneously (i.e. one experiment). At least two independent experiments were performed for each figure, and the data shown are representative. The editing activity varied between editosome preparations, but the relative efficiency of the constructs was always consistent. The abundance of accurately edited and cleavage product for each construct tested was initially calculated as the percentage of total input RNA and then normalized to the abundance of the corresponding product by the parental Pair-1 substrate. Mean and error bars were plotted on a linear scale.

Annealing and Photo-cross-linking Assays

The efficiency of pre-mRNA/gRNA annealing was scored in native gels. ES2-radiolabeled mRNA(~10 fmol) and gRNA (1.2 pmol) were pre-annealed in a 2-µl mixture for 10 min at 37 °C followed by 1h at 26 °C, as for editing assays. 20-µl mixtures were completed with 10 mM MRB buffer (see above) containing xylene cyanol and bromphenol blue, incubated for an additional 10 min at 26 °C, loaded directly onto a 6% native PAGE, and resolved at 180 V for 6 h at 4 °C. 0.5 X Tris-borate EDTA buffer and 1 mM $MgCl_2$ were used in both the gel and running buffer. The photo-cross-linking assays were performed using pre-annealed RNA pairs and under editing reaction conditions (but without nucleotides) as recently reported (94). Also, coimmunoprecipitation and competition analyses were carried out as described (94). All competitor transcripts were supplemented to mixtures and incubated for extra 10 min at 26 °C to allow annealing (i.e. of homologous competitor with free cognate gRNA) before the addition of complexes and irradiation.

Results

Analysis of the Natural A6 Pre-mRNA Features Proximal to ES2 for Full-round Insertion

Features in the RNA substrate that are specifically recognized during full-round editing are not fully defined in trypanosomes. These may include the native pre-mRNA

sequence, purine richness, and higher-order structure of the pre-mRNA/gRNA bimolecular substrate. To address this question we characterized the proximal features of an A6 RNA pair (Pair-1) for ES2 insertion that uses a 45-nt pre-mRNA fragment (Fig. 16A; Ref. 41). We have shown that minimal 43–45-nt pre-mRNA substrates support efficient fullround insertion in the A6, Cyb, and RPS12 systems (91). For simplicity, we will refer to the upstream and downstream duplexes (relative to the pre-mRNA) flanking the internal loop containing ES2 as "left" and "right" helices. The terminal 5' extensions of Pair-1 will be termed pre-mRNA and gRNA protrusions, respectively (Fig. 16A).

We first analyzed the left helix of Pair-1. Our previous studies showed that virtually all natural 5' purines in the pre-mRNA protrusion could be substituted by pyrimidines (91). It was also reported that natural pre-mRNA sequence beyond the residues forming the right duplex were dispensable for ES2 insertion (83). To assess the importance of the natural pre-mRNA sequence in the left helix and the requirement for the pre-mRNA protrusion, we designed Pair-2 containing an artificial 15-bp blunt-ended left duplex (Fig. 16A). Furthermore, the pre-mRNA/gRNA-paired residues in this duplex were flipped to alternate all purines and pyrimidines (except the first two residues needed for T7 *in vitro* transcription). Pair-2 supported insertion at a level comparable with the parental Pair-1 (Figs. 16, B and C), thus showing that neither the natural pre-mRNA sequence nor purine richness in the left duplex or the pre-mRNA protrusion is required for efficient insertion. Note that to preserve both a pre-mRNA functional length (91) and its purine content in Pair-1 (77.8%), all 25 upstream purines in the parental substrate were moved to the 3' end of Pair-2 (75% purines).

To further analyze the functional length of the left helix, we tested Pair-2 derivatives (Fig. 16A) containing 12-bp (Pair-3) or 11-bp (Pair-4) helices with predicted stabilities similar to the parental duplex (data not shown). Interestingly, Pair-3 was edited less efficiently than Pair-1. This may reflect a partial influence of nucleotide composition of the helix. Notably, the artificial 11-bp upstream duplex in Pair-4, which represents ~one turn of helical RNA (237), efficiently replaced the complete 5' region of the parental Pair-1. Because the minimal length of the natural A6 pre-mRNA for efficient full-round ES2 insertion is ~43–45 nt (91), we trimmed the rightward region of Pair-4 to generate Pair-5 (45-nt pre-mRNA; Fig. 17A). This substrate, with reduced purine content (71%), supported less accurate editing (i.e. 3U addition) than the parental Pair-1 and accumulated inaccurate insertion by 1U addition (Fig. 17B). However, relocation of the protruding 3' purines of Pair-5 to the 5' end (Pair-6; 71% purines) re-established 3U insertion nearly to the level of the parental Pair-1 (lane 3). Although the latter 45-nt constructs imply that a short protrusion may be more stimulatory 5' than 3' to ES2, alternative structural reasons are also feasible. To determine whether or not the pre-mRNA extension in Pair-6 must be purine-rich, we substituted most of the protruding purines by pyrimidines (Pair-7). Interestingly, Pair-7 was about as efficient as Pair-1 despite its relatively low (49%) purine content (Fig. 17C, compare the first and last lanes).

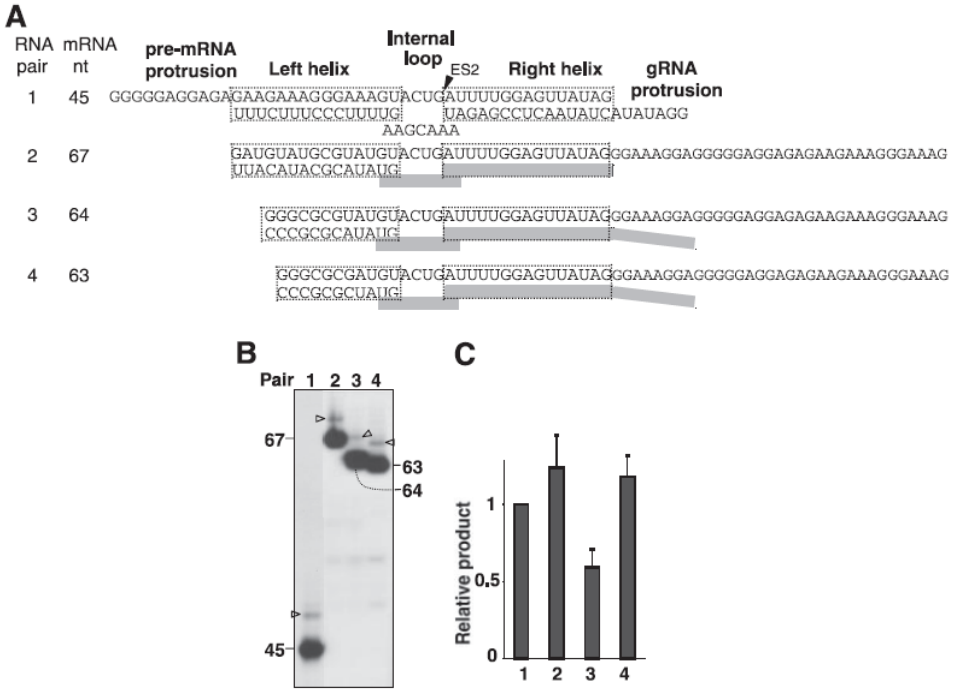

Figure 16. Analysis of the left helix in the parental A6 pre-mRNA/gRNA substrate for full-round insertion at ES2 (Pair-1) (41). A, the starting Pair-1 substrate and nomenclature of the analyzed regions. Derivative constructs are aligned. The gRNA gA6[14]USD-3A (83) in Pair-1 (in blue) and the ES2 (arrowhead) are indicated. The predicted stability (3.0 Mfol, M. Zuker program (238)) of left and right duplexes (boxed) of Pair-1 are -20.5 and -21.5 kcal/mol. The based composition of the derivative duplexes was adjusted to conserve a similar ΔG° as in Pair-1. In the left columns the RNA pair assigned number and pre-mRNA size are indicated. Parental unmodified gRNA sequences are depicted by filled boxes. B, full-round insertion assays using 3'-end-labeled pre-mRNA constructs. Accurate insertion by addition of three Us is indicated by an arrowhead. The pre-mRNA size and assigned pair number are indicated. C, plots of relative accumulation of accurately edited product using 3'-end radiolabeled pre-mRNA (see "Experimental Procedures").

Together, the above results indicate that neither the natural polypurine run nor the overall purine richness of A6 pre-mRNA is a critical determinant of ES2 insertion *in vitro*. Combined with our previous observations (91), these features upstream of the internal loop is sequence-independent. Furthermore, a single turn of helical RNA was sufficient for full-round insertion. Additional work will be required to test whether smaller and less stable left helixes are efficient.

We then examined the features of the right helix of Pair-1 required for editing. To this end we used variants of Pair-6 (Fig. 17A) containing 14-bp (Pair-8) and 13-bp (Pair-9) right duplexes with a predicted stability comparable with the parental helix (data not shown). Both Pair-8 and Pair-9 supported editing nearly as efficiently as the parental Pair-1 (Fig. 17C). We then tested derivatives bearing either 12-bp (Pair-10) or 11-bp (Pair-11) right helices.

Notably, whereas the base composition of the 3' duplex significantly deviated from the parental helix, these substrates were appreciably more efficient than Pair-1 (Figs. 17, D and E). It is also worth noting that the predicted stability the right helix in Pair-11 is ~10% lower than in Pair-1 (see the legend to Fig. 17).

Finally, we generated Pair-12 bearing a blunt-ended right duplex of 11 bp, and a change of the 5'-terminal two gRNA residues to facilitate *in vitro* transcription starting with G (215). This last construct was less efficient than Pair-11. It is possible that a branch structure downstream of ES2 favors insertion editing. Consistent with this notion, a derivative of Pair-10 with a right blunt-ended duplex was also less efficient (data not shown). Altogether, these data show that substrates with one turn of helical

RNA at both the left and right sides of ES2 support efficient full-round U insertion and that the natural sequence of the minimal (~43–45 nt) pre-mRNA outside the internal loop is not required. This suggests that the basic editing complex recognitions flanking the internal loop involve sequence-independent features of the pre-mRNA/gRNA pair.

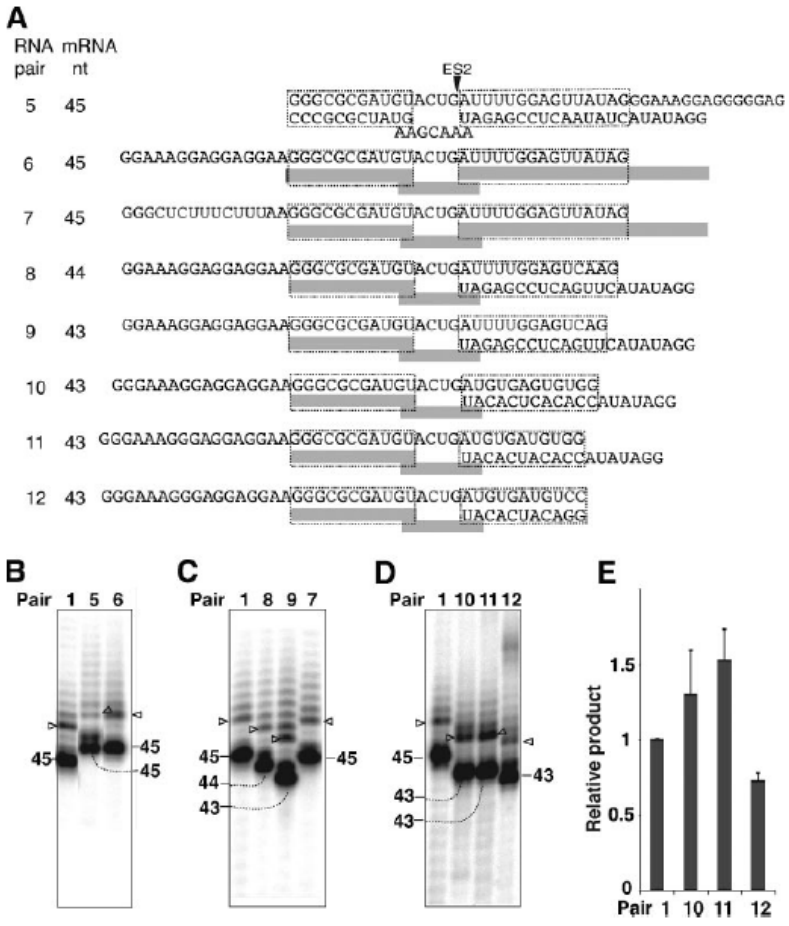

Figure 17. Analysis of the right helix. A, the starting Pair-5 and derivatives are aligned. All labeling is as in Fig. 1A. B–D, full-round insertion assays are as in Fig. 1B. The derivative duplexes conserved a similar predicted ΔG°, except for the right helix in Pair-11 and Pair-12 that dropped to -18.3 kcal/mol. The pre-mRNA size in these pairs was adjusted to a minimum of 43 nt (41), as the 3' end was truncated. Some sequence-dependent gel migration differences were observed (e.g. the 3' end of the last two 43-nt pre-mRNAs differ by two residues). E, plot as in Fig. 16C.

Effect of Ribose 2' Substitutions on Full-round Insertion

We analyzed the contribution of ribose 2'-hydroxyl groups to substrate recognition by editing complexes, by incorporating 2'-deoxy substitutions and other 2' modifications in and around ES2 in the parental Pair-1 (summarized in the figure on p. 105).

Pre-mRNA Residues Upstream of the Internal Loop

We first analyzed the pre-mRNA 5' region (Fig. 18A) using derivatives containing 11 deoxynucleotides that replaced either the entire pre-mRNA protrusion (Pair-13) or most bases in the left helix (Pair-14). Interestingly, both constructs supported insertion at about half the level of the parental Pair-1 (Fig. 18B, upper panel). This decrease in editing was largely paralleled by a reduction in ES2 cleavage (Fig. 18, B, lower panel, and C). Importantly, pre-mRNA cleavage was scored in absence of RNA ligase activity using editing complexes that were pretreated with 10 mM Ppi (Ref. (73); see "Experimental Procedures"). These two constructs showed that 2'-hydroxyl groups in the pre-mRNA protrusion and the left helix partially stimulate insertion. Our previous work showed that truncation of the protrusion in the minimal 45-nt A6 pre-mRNA (see 34-nt RNA 8 in Cifuentes-Rojas et al. (91)) abrogates full-round insertion. This suggests that editing complexes make sufficient contacts with the all-DNA protrusion to support an appreciable insertion level, and thus, the 2'-hydroxyls of the pre-mRNA protrusion are significantly stimulatory but not essential. Furthermore, the left RNA/DNA heteroduplex in Pair-14 should adopt a shape that is more similar to the

A-form (of dsRNA) than the mainly reflect a requirement for 2'-hydroxyls, although shape dependent recognitions may also be important. We then tested the effect of a single 2'-deoxynucleoside substitution adjoining the scissile bond (Pair-15). Notably, this 2'-H abolished both insertion and cleavage (Figs. 18, B and C). This may reflect a direct role of the 2'-OH at ES2 in catalysis and/or site recognition or an indirect role due to a change in the sugar pucker (from ribose C3'-endo to deoxyribose C2'-endo conformation (237).

To address these two possibilities we tested other ribose 2' modifications such as 2'-O-methyl (-OCH_3; Pair-16) that favor the RNA-like sugar pucker but cannot act as a hydrogen bond donor (57). 2'-OCH_3 supported neither insertion nor cleavage (Figs. 18, B and C), but it is conceivable that the added bulk of this group, compared with a 2'-OH, caused steric hindrance. We then tested the smaller 2'-fluorine (-F) modification (Pair-17), which should also promote RNA-like C3'-endo conformation even more than 2'-OCH_3 and is highly unlikely to accept a proton (236). This substituent also completely inhibited cleavage (Fig. 18, C and D; see the legend). Unfortunately, a 2'-NH_2 modified guanosine at ES2 is not available (Dharmacon).

Based on these results, it is conceivable that formation of a hydrogen bond by the ribose 2'-OH group at ES2 is required for insertion. The ribose 2'-hydroxyl group at ES2 could mediate either catalysis at pre-mRNA cleavage, editing site recognition, or editing complex assembly onto the substrate.

To attempt distinguishing between these possibilities, we modified our recently developed photo-crosslinking assay to analyze direct editing complex contacts at ES1 in

A6 pre-mRNA (94). To similarly assay ES2, we placed a single ^{32}P-labeled 4-thioU immediately 3' of the scissile bond. Pairing of this residue with a guiding adenosine should extend the right helix by 1 base pair (Pair-15'; Fig. 18E). Notably, the thiolated ES2 supported protein crosslinking interactions that both are similar to those reported for ES1 (94) and specifically co-purify and co-immunoprecipitate with editing complexes (see Fig. 22). Substrates with either a 2'-OH (Pair-1') or the inhibitory 2'-H modification (Pair-15) at ES2 exhibited identical cross-linking patterns (Fig. 18E), suggesting a similar editing complex association with both the 2'-H-substituted and unmodified ES2. Thus, the single 2'-deoxy substitution at ES2 does not seem to prevent editing complex interactions at ES2, adjoining the scissile bond may play a role in catalysis.

Pre-mRNA Residues in the Right Helix

We then examined the pre-mRNA residues in Pair-1 that are part of a 15-bp right duplex. Three deoxynucleotide patches were initially compared, 10, 14, and 15 nt long (Pairs 18, 19, and 20, respectively; Fig. 18A). Note that these pre-mRNAs contain a 3'-most ribonucleoside to allow end-radiolabeling with T4 RNA ligase (see "Materials and methods" (236). Interestingly, these ribose substitutions decreased both insertion and cleavage compared with Pair-1 (Figs. 18, F and G). This inhibition increased with the extent of deoxy substitution. Notably, Pair-20 with all upstream pre-mRNA residues modified was most inhibited.

To determine the importance of the 2'-hydroxyl immediately 3' of the cleavage site, the singly substituted Pair-21 was tested. This substrate was about half inhibited in both editing and cleavage assays (Figs. 18, F and G). Combined, these observations indicate that the 2'-hydroxyls just 3' of ES2 and further downstream in the duplexed pre-mRNA are significantly stimulatory. These deoxy substitutions did negatively impact insertion, primarily at pre-mRNA cleavage.

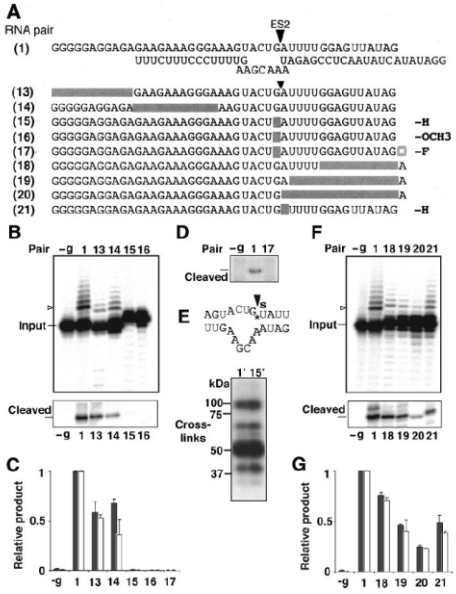

Figure 18. Ribose 2'-substitution of upstream and downstream residues in the pre-mRNA of Pair-1. A, Pair-1 and derivatives are aligned (pre-mRNA strand is shown). Single or multiple 2'-substitutions are indicated in boxes. The modifications are 2'-deoxy (-H), 2'methoxy (-OCH_3) or 2'-fluorine (-F). The latter substrate was manufactured with a 3'deoxynucleoside (empty box). All other labeling is as in Fig. 16. B, insertion and cleavage (upper and lower panels, respectively) of 5'-substituted pre-mRNAs. A control lane devoid of gRNA (-g) was included. C and G, relative accumulation of accurately edited (black bars) and ES2 cleavage (white bars) of 3'-end labeled pre-mRNA. D, cleavage of 2'-fluorine modified, 5'-end labeled mRNA (see "Experimental Procedures"). Precise cleavage at ES2 was confirmed using ribonuclease T1 and hydroxyl partial pre-mRNA digestions (not shown). E, editosome photocross-linking with pre-mRNAs containing a single ^{32}P and 4-ThioU at ES2 of Pair-1 and Pair-15. Diagram indicates the labeled bond (*) and thiolated U (s). The right duplex was extended by one base pair between Thio-U and a guiding adenosine. The size of molecular markers is indicated in kDa. F, insertion and cleavage assay of 3'-substituted pre-mRNAs.

gRNA Residues and Duplexes Flanking the Internal Loop

Apart from the critical ribose 2'-OH at the editing site, most pre-mRNA 2'-hydroxyl groups tested were stimulatory but not essential for ES2 insertion. We then examined the effect of proximal gRNA substitutions and DNA duplexes (Fig. 19A). A 10-deoxynucleotide patch on the gRNA strand at either side of the internal loop (Pair-22 and Pair-24, respectively) had a slight negative effect on insertion (Fig. 19B) comparable with that observed with corresponding pre-mRNA patches (Pair-14 and Pair-18). In contrast, DNA duplexes formed by the complementary patches at left (Pair-23) or right (Pair-25) of ES2 were more inhibitory, particularly the Pair-25 (Fig. 19B). In both cases, insertion and pre-mRNA cleavage at ES2 were similarly inhibited (Fig. 19D).

Interestingly, the right DNA duplex also significantly affected the scissile-bond selectivity. That is, the pre-mRNA of Pair-25 was cleaved at several residues flanking ES2; the upstream cuts are in the loop, whereas the downstream cuts are in the right duplex. All these cleavages were gRNA-dependent (not shown). We assessed whether the inhibitory DNA duplexes affected the editing complex photo-cross-linking interactions with ES2. Interestingly, thiolated versions of these substrates (Pair-23' and Pair-25') reduced the level of all cross-linking subunits (Fig. 19E). This effect was particularly severe with Pair-25'. However, because protein-RNA cross-linking can be quite sensitive to conformational changes, inhibition of cross-linking activity may reflect local structural changes of the editing site rather than reduced editing complex assembly onto the RNA substrate. A native gel analysis of DNA duplex-containing pairs and Pair-1 confirmed complete annealing of these substrates (Fig. 19F).

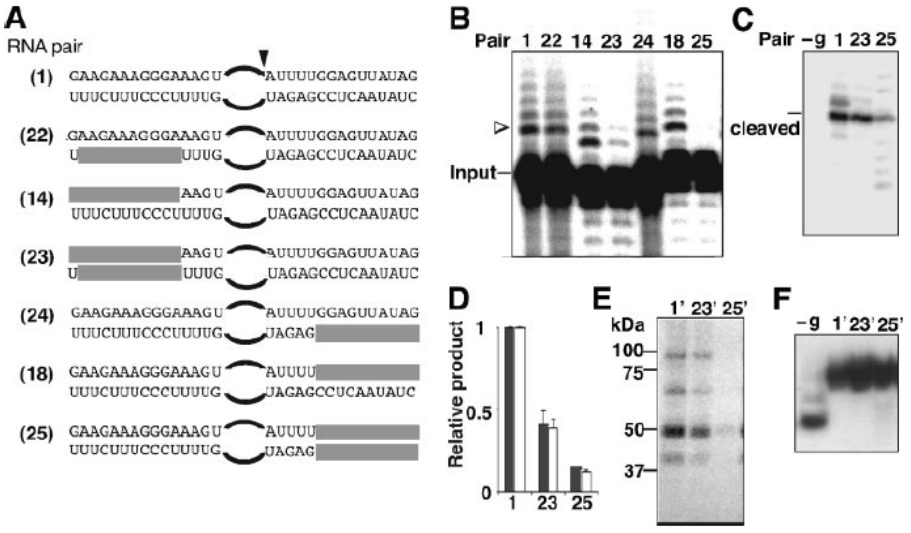

Figure 19. Ribose 2' substitution of Pair-1 pre-mRNA and gRNA strands. A, pair-1 and derivatives are aligned. Internal loop residues around ES2 are depicted as curved lines with an arrowhead pointing to ES2. Deoxynucleoside substituted residues are boxed. B, insertion assay. C, cleavage assay. The inaccurate cleavages with Pair-25 are also gRNA dependent (data not shown). Plots (D) and photo-cross-linking assays (E) are as in Fig. 18. F, annealing assays of RNA pairs used in E. Control lanes without gRNA (-g) are indicated.

Combined, the parallel inhibition of insertion, cleavage, and cross-linking activities at ES2 suggest that proximal DNA duplexes negatively impact a productive interaction of editing complexes with the substrate and, thereby, catalysis.

Internal Loop Residues

Finally, we examined the importance of 2'-hydroxyl groups in the internal-loop residues containing ES2 (Fig. 20A). Deoxy substitutions in the three residues 5' of the guanosine at ES2 (Pair-26) had virtually no effect on either insertion or cleavage (Fig. 20, B and C). In contrast, 2'-H substitution in all seven loop gRNA nucleotides (Pair-27) significantly reduced editing and cleavage (Fig. 5, C and D). Notably, modification of both strands of the internal loop (Pair-28) further inhibited both editing and cleavage. Furthermore, editing complex cross-linking at ES2 was also moderately and strongly reduced in the corresponding Pair-27' and Pair-28' substrates, respectively (Fig. 20E). A native gel analysis confirmed the complete annealing of these pairs (Fig. 20F).

Together, these observations indicate that several hydroxyl groups in the internal loop are relevant determinants of insertion that largely influence both the efficiency of pre-mRNA cleavage and cross-linking by editing complexes. Interestingly, hydroxyls on the gRNA strand appeared to compensate for deoxy substitutions on the pre-mRNA strand but not vice versa. That is, one or more 2'-hydroxyls in the gRNA internal loop residues significantly stimulate *in trans* pre-mRNA cleavage and/or site recognition by editing complexes. The analyses in Figs. 18–20 are summarized in Fig. 21.

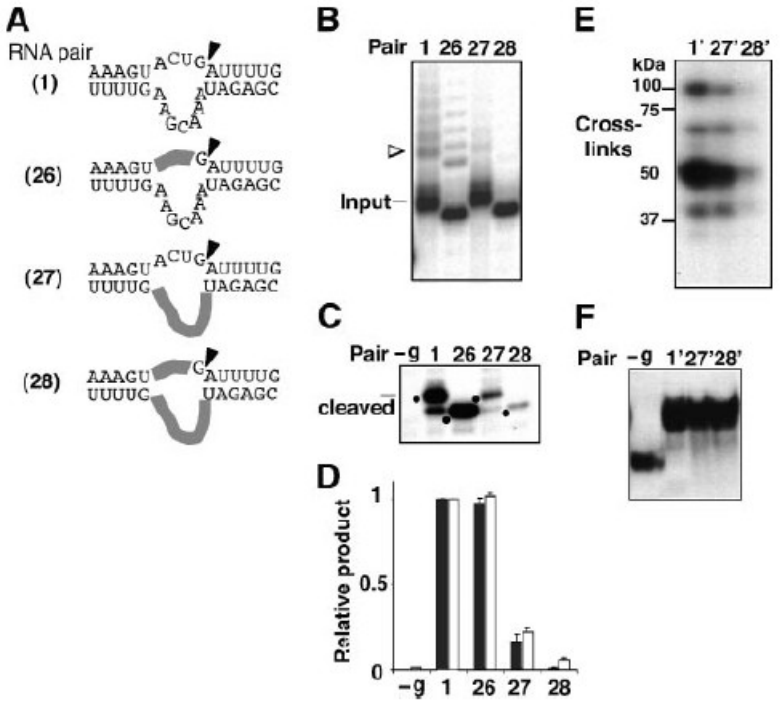

Figure 20. Ribose 2' substitution of internal loop residues around ES2 in Pair-1. A, Pair-1 and derivatives. B, insertion assay. C, cleavage assay. ES2-cleavage products are marked with filled dots. Note that the offset gel mobility reflects the use of the same pre-mRNA in Pairs 1 and 27 as well as in Pairs 26 and 28. D, plots and E, photo-cross-linking assays are as in Fig. 18. F, annealing assays of RNA pairs used in D. Control lanes without gRNA are indicated.

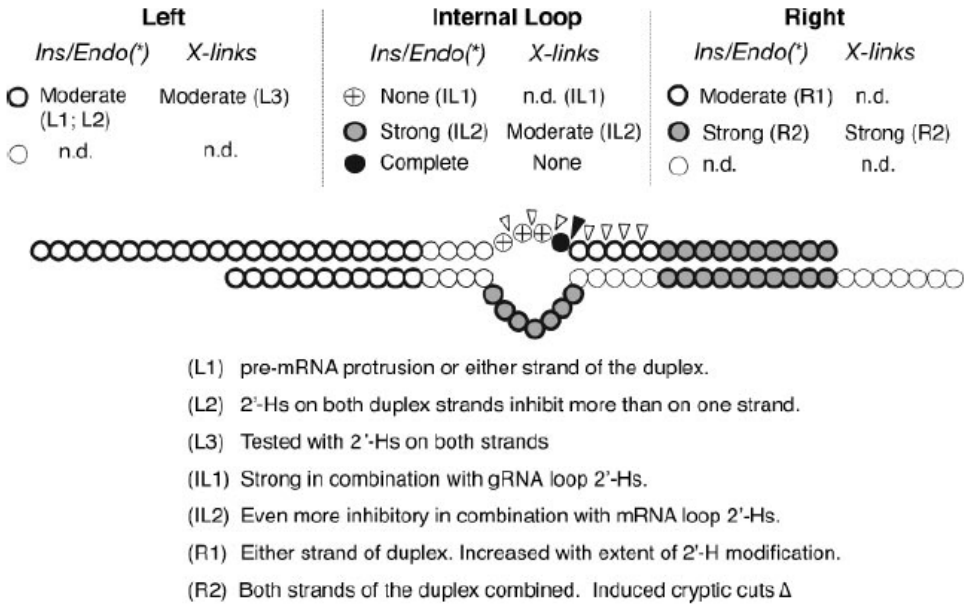

Figure 21. Summary of ribose 2'-deoxy substitutions tested on Pair-1 for full-round insertion at ES2. Upper panel, indicates the level of inhibition for the left and right sides of the internal loop and within the loop (each region is separated by a vertical line). Full-round insertion and pre-mRNA cleavage (Ins/Endo) are on the left, and editosomes cross-linking (X-links) are at the right. The asterisk indicates that insertion and cleavage are affected at comparable levels. Circle types representing the level of inhibition: thick line, moderate; thin line, not determined (n.d.); gray (in addition to thick line), strong; with a pattern, no effect; black, complete. Parentheses indicate clarification notes (lower panel). A filled arrowhead points to the natural ES2 for full-round insertion. dsDNA within the right duplex induced cryptic cuts at several residues (open arrowheads) flanking the editing site. Middle panel, diagram of Pair-1 with individual residues shown as circles. Lower panel, explanatory notes on the effect of the indicated modifications.

Specificity of Cross-linking Interactions at the Insertion ES2

Several observations indicate that the protein-RNA photo-crosslinking interactions at the insertion ES2 represent direct contacts of editing complexes with the substrate. For example, the cross-links are specifically co-immunoprecipitated by antibodies raised against protein subunits of the complex (Fig. 22A and data not shown). Also, native complexes purified by either two or three steps of consecutive ion-exchange chromatography exhibit comparable cross-linking (Fig. 22B) and editing activities at ES2 in the minimal substrate analyzed (91). Thus, although the latter preparation has a simpler protein composition (Fig. 22C), the two preparations of editing complexes appear functionally equivalent with the substrate analyzed. Furthermore, the presence of representative subunits (Fig. 22D and data not shown) as well as all critical catalytic activities including editing endonucleases suggest that the functional and protein composition of our complexes is similar to that reported by other groups (54,55).

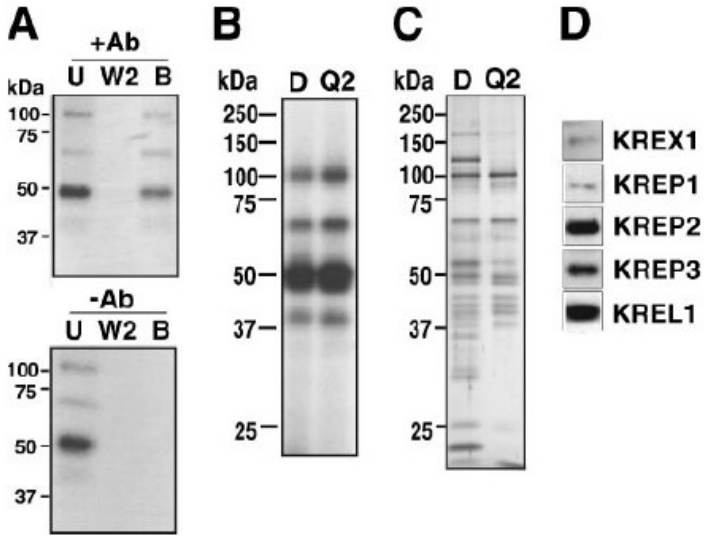

Figure 22. The RNA-protein cross-linking interactions at ES2 co-immunoprecipitate and co-purify with native editing complexes obtained by sequential steps of chromatography. A, an immunoprecipitation assay using anti-KREPA3 monoclonal antibodies (Ab), including the unbound (U), second wash (W2), and bound immunoprecipitated (B) fractions (upper panel). A parallel mock assay without antibodies (-Ab) is also shown (lower panel). B, the cross-linking pattern is conserved in the peak fractions of consecutive chromatography steps; D and Q2 represent two-step and three-step purification protocols, respectively (see "Experimental Procedures"). C, silver-staining of the fractions in B. D, Western blot analysis of the Q2 fraction with available antibodies against five different stably-bound subunits of editing complexes.

Together, these results indicate that the cross-linking interactions at ES2 are specific to editing complexes. This is consistent with our recently reported observations of cross-linking interactions by editing complexes at the deletion site ES1 in a similar A6 substrate (94). We analyzed the specificity of the cross-linking and pre-mRNA cleavage activities at ES2 in competition analyses with homologous and heterologous transcripts (Figs. 23, A and B). In both cases the bimolecular A6 substrate was readily competed out by a 5- and 10-fold excess of homologous pre-mRNA that can hybridize with free cognate gRNA (~120 times the pre-mRNA concentration in the standard mixture; see "Materials and methods"). In contrast, heterologous transcripts including another pre-mRNA, tRNA, and a non-complementary gRNA were partially or not inhibitory at greater (10- and 25-fold) excess. Finally, further addition of cognate gRNA was not inhibitory. Thus, both cross-linking and cleavage activities of editing complexes at ES2 exhibit specificity for the hybrid substrate. These observations are also consistent with our reported preferential interactions of editing complexes at ES1 (94). Interestingly, the highly structured tRNA competitor partially affected both cross-linking and editing activities more at ES2 than at ES1 (Ref. (94); Fig. 23). The similar gel mobility of mayor cross-links at ES2 and ES1 (Fig. 24) suggested that the same subunits of editing complexes make these contacts. However, the cross-linking efficiency at ES1 is significantly greater than at ES2, consistent with the former substrate supporting a much higher level of editing *in vitro* (51).

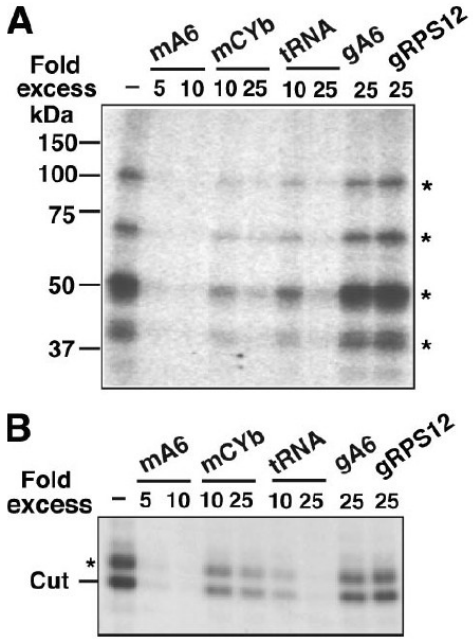

Figure 23. Competition analysis of cross-linking and pre-mRNA endonuclease activities of editing complexes. Cross-linking (A) and cleavage (B) assays with or without a 5-, and 10-molar excess of homologous A6 pre-mRNA (mA6) or 10- and 25-fold excess of heterologous Cyb mRNA (mCYb), tRNA, and non-complementary gRNA gRPS12. Extra cognate gRNA (gA6) was also supplemented to the standard mixture containing 120-fold excess of this transcript (see "Experimental Procedures"). The cleaved pre-mRNA piece (Cut) is mixed with a 1-nt 3 – extended homologue fragment (asterisk) derived from T7 RNA polymerase *in vitro* transcription.

Because the cross-linking patterns at ES2 and ES1 are similar, we asked whether a transcript that does not undergo editing can cross-link with editing complexes. To this end we tested gRNA D33, which does not inhibit ES1 cross-linking at a ~200-fold excess and exhibits virtually no predicted structure (94). Surprisingly, such a transcript containing a single photoreactive 4-thioU supported a level and pattern of cross-linking comparable with that of ES2. Nevertheless, as expected, an excess of the homologous competitor inhibited all cross-linking by thiolated D33, whereas the same (20–40-fold) or greater (200-fold) excess of competitor had no effect on the cross-linking interactions at ES2 and ES1 (Fig. 24 and data not shown).

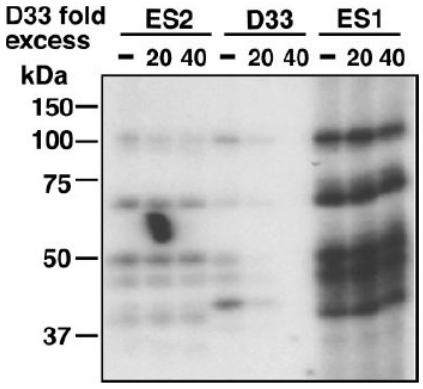

Figure 24. Competition analysis of cross-linking interactions with bimolecular editing substrates and a non-edited transcript. Lanes 1–3, 2–5, and 7–9 show interactions with ES2, ES1, and D33 photo-reactive substrates, respectively. The D33 competitor was used at the indicated molar excess.

Overall, these observations suggest that the cross-linking subunits of editing complexes can make transient nonspecific contacts with RNA; however, the similar interactions at ES2 and ES1 are significantly more stable.

Discussion

The molecular basis of substrate recognition by editing complexes and the regulation of RNA editing in the single mitochondrion of trypanosomes are still poorly understood. The purpose of these studies was to dissect functional substrate determinants proximal to a site for full-round U insertion catalyzed by purified editing complexes. Combined, these observations and our previous study (94) have important implications on the mechanisms of substrate recognition by editing complexes. First, the overall recognition of the insertion substrate outside the internal loop is sequence independent. This notion is consistent with structural probing studies suggesting that related secondary structures of different mRNA/gRNA pairs may be important for editosome recognition (234).

Our analysis of a minimal A6 RNA pair for full-round insertion at ES2 (91) showed that the sequence and base composition of the parental helices flanking the editing site, including the pre-mRNA purine-richness (93%) in the left duplex, are not required for efficient editing. We had recently shown that the all-purine pre-mRNA protrusion could be replaced by a pyrimidine-rich stretch (91). Although natural 5' poly-purine runs in the A6 pre-mRNA are dispensable for the basic insertion reaction *in*

vitro, it is conceivable that these structures are specifically recognized *in vivo* by factors to promote nucleation with the complementary uridylate tail of gRNAs or stabilization of the duplex. In line with this notion, an editing complex subunit (KREPA4) was recently reported to exhibit binding specificity for a gRNA 3' U-tail (37) (229). Furthermore, a proposed accessory factor (REAP-1) preferentially binds to purine-rich transcripts such as pre-mRNAs (239). Notably, substrates with completely artificial 11-bp duplexes (i.e. one helical turn of RNA) flanking ES2 support efficient insertion (e.g. Pair-11 and Pair-12; Fig. 17).

We speculate that one or both duplexes flanking ES2 is recognized by KREN2, an endonuclease subunit that specifically serves in insertion (59) and bears an RNase III-like domain and one double-stranded RNA binding motif (46,54). Interestingly, the 11-bp artificial helices in our substrates may be minimal in length, as structural studies of highly conserved double-stranded RNA binding motifs in other systems indicate that these proteins typically interact with 16-bp (~1.5 helical turns of dsRNA; Ref. (240)). Moreover, a recent study proposed that the smallest dsRNA substrate for the single double-stranded RNA binding motif in a bacterial Rnases III is 11 base pairs (149). KREN2 may dimerize (24) like other class 1 RNase III enzymes (152), so that each double-stranded RNA binding motif could contact one of the 11-bp helices flanking ES2 in our constructs. Typical double-stranded RNA binding motifs specifically bind the A-form of dsRNA through interactions that are adapted to the shape of the helix, are sequence-independent, and primarily involve hydrogen bonds with ribose 2'-hydroxyls (241). This is consistent with the observed inhibition of both U insertion and editing

complex photo-cross-linking interactions at ES2 by the presence of (B-form) dsDNA in the flanking helices (Fig. 19B). Also in line with this notion, 2' deoxy substitutions in one strand caused more moderate negative effects (Figs. 19, B and E) possibly because DNA/RNA heteroduplexes retain half of the hydroxyls and conserve more the RNA than the DNA helical shape (242). An 11-bp DNA duplex 5 bp downstream of ES2 (Pair-25) markedly reduced accurate cleavage and, instead, stimulated low-level cryptic pre-mRNA cuts near ES2 (Fig. 19C). Such an effect on scissile-bond selection suggests that the modified riboses in the right duplex help position KREN2 to precisely cleave the bond at the single-/double-strand junction (ES2). Interestingly, the significant cleavage inhibition correlates with a dramatic reduction of editing complex cross-linking at the same site (Fig. 19E). Similarly, the presence of helical DNA 4 bp upstream the internal loop moderately inhibited both cleavage and RNA-protein interactions at ES2.

Together, these data suggest that the editing complex makes important contacts with both helices flanking ES2. Such contacts with the downstream anchor duplex appear particularly relevant for efficient and accurate pre-mRNA cleavage as well as for cross-linking interactions at the editing site. 2'–hydroxyl groups in the pre-mRNA protrusion (i.e. Pair-13; Fig. 18A) also contribute to the insertion reaction but more moderately. However, a construct with the combined modifications of Pairs 13 and 14 significantly inhibited cleavage (data not shown) and, thus, further emphasize the impact of the upstream pre-mRNA structure on this editing step. Single ribose 2' substitutions at ES2 including 2'-H, 2'-OCH_3, and 2'-F showed that the 2'-hydroxyl at this site is critical for either pre-mRNA cleavage or prior ES2 recognition. Interestingly, ES2 with

either a 2'-OH or 2'-H supported the same pattern of editing complex cross-linking interactions at the editing site. Because protein-RNA photo-cross-linking is sensitive to substrate conformational changes, we suspect that most protein interactions involved in ES2 recognition are unaffected at the 2'-H modified site.

A potential role of the 2'-hydroxyl at ES2 in catalysis could involve 1) formation of a critical hydrogen bond with the KREN2 endonuclease, 2) coordination of a divalent cation cofactor, or 3) direct nucleophilic attack on the scissile bond. Editing complex-catalyzed cleavage of pre-mRNA containing a single ^{32}P at ES2 produces a 5'-labeled downstream fragment (data not shown) rather than a 3'-labeled upstream fragment as would be expected if the 2'-OH group forms a 2',3' cyclic phosphate upon cleavage, as occurs with RNase A (81). This is consistent with previous RNase mapping that deduced the 5'-P and 3'-OH nature of the termini produced at ESs (48,50) and also is in line with an RNase III Mg^{2+}-dependent type of processing (48,50). Further work will be needed to dissect the precise role/s of the 2'-hydroxyl group at ES2 in pre-mRNA cleavage and/or editing site recognition. Our deoxy substitutions within the ES2 internal loop were also quite informative. Interestingly, substitution of the three residues 5' of the guanosine (bearing a critical 2'-OH) at ES2 affected neither cleavage nor U-insertion; however, substitution of all apposing loop gRNA residues strongly inhibited both insertion at pre-mRNA cleavage and cross-linking activities at ES2. These results were unanticipated as they reveal that the gRNA strand in the internal loop is an important determinant of full-round U-insertion at the level of pre-mRNA cleavage and suggest that the proximal pre-mRNA 2'-OH groups upstream of the guanosine at ES2

are less relevant for substrate cleavage. Because the relative level of insertion and cleavage were similar, these 2'-deoxy modifications appear to have little or no effect on either U addition (by the terminal U-transferase, KRET2) or RNA ligation (by KREL2) in the insertion cycle. It is 115mmune115nsc that the combined pre-mRNA/gRNA substitutions in the loop were more inhibitory (in all assays tested) than gRNA modifications alone because pre-mRNA substitutions had no effect. It is possible that the pre-mRNA deoxynucleotides facilitate a conformation of the substituted gRNA loop that is particularly inhibitory. Several editing complex subunits contain conserved motifs that may bind single-stranded RNA around an ES. For example, KREPA1 (also termed band II, LC-1; Refs.(47,54,203)) has an oligonucleotide/oligosaccharide binding (OB)-fold that could bind single-stranded RNA in interactions independent of ribose 2'-hydroxyls (243). KREPA1 was speculated to have a role in recognition of the editing substrate and possible coordination of an insertion cycle (54). Consistent with this concept, we recently proposed that the related OB-fold containing KREPA2 (band III, LC-4) directly binds a site for full-round U deletion (94). Additional related subunits of the editing complex were also proposed to conserve an OB-fold. Three subunits including KREPA1 and KREPA2 also bear C_2H_2 zinc-finger domains that could potentially bind single-stranded RNA (241,244). Furthermore, the RNase III-like insertion endonuclease KREN2 is expected to cleave single-stranded RNA at a single-/double-strand junction, unlike typical RNase III enzymes that cleave dsRNA. KREN2 has one double-stranded RNA binding motif, one RNase III domain, and also a U1-like

zinc finger (54). Perhaps one or more of these motifs specifically interact with internal loop determinants involved in scissile-bond selection at insertion sites.

All ribose 2'-deoxy substitutions tested in this study are summarized in Fig. 21. Notably, proximal changes that significantly decreased pre-mRNA cleavage were also associated with a parallel inhibition of editing complex cross-linking at ES2. Thus, the cross-linking assay we introduced here not only revealed for the first time direct editing complex interactions at an ES for full-round insertion but also can help define ribose 2'-OH groups in cis (pre-mRNA) or in trans (gRNA) that affect the efficiency of both cleavage catalysis and photo-cross-linking at sites targeted by editing complexes.

Work is in progress to identify the photo-cross-linking subunits; however, due to the similar size and gel mobility of several subunits, identification of the cross-linking proteins is not straightforward. We are currently combining the use of mass spectroscopy techniques and epitope-tagging to produce confirmatory gel-shifts of candidate subunits. Interestingly, the cross-linking pattern at the insertion ES2 and deletion ES1 (94) in A6 model substrates are similar. Whether or not they play a role in the distinction of editing sites is uncertain, although the cross-linking efficiency at ES1 is significantly greater than at ES2. Notably, the former is the most efficient model ES known for *in vitro* editing (51).

Our competition analyses showed a similar response of both cross-linking and pre-mRNA cleavage activities and suggest a preferential association of editing complexes with the A6 substrate. These results also implied a functional relevance of the cross-linking interactions. Surprisingly, a transcript (D33) that does not interfere

with cross-linking at an ~200-fold excess (94) cross-linked in its thiolated version with editing complexes. However, as expected, these cross-links were very sensitive to low concentrations of unlabeled D33 (competitor), whereas the similar ES2 and ES1 cross-links were resistant to the same or greater concentrations of D33. This suggests that most if not all crosslinking subunits can make transient nonspecific contacts with RNA, which may be detected by the sensitive 4-thioU photoreagent. In contrast, the associations at ES2 and ES1 are relatively stable. These stabilized interactions, however, are sequence-independent as they are conserved at both ESs tested and may reflect recognition of helical irregularities (e.g. potential ESs) in hybrid substrates. Our laboratory is currently exploring these possibilities.

CHAPTER IV

DETERMINANTS FOR ASSOCIATION AND gRNA-DIRECTED ENDONUCLEASE CLEAVAGE BY PURIFIED RNA EDITING COMPLEXES FROM *TRYPANOSOMA BRUCEI**

Summary

U-insertion/deletion RNA editing in the single mitochondrion of ancient kinetoplastids is a unique mRNA maturation process needed for translation. Multi-subunit editing complexes recognize many pre-mRNA sites and modify them via cycles of three catalytic steps: guide-RNA (gRNA) mediated cleavage, insertion or deletion of uridylates at the 3' terminus of the upstream cleaved piece, and ligation of the two mRNA pieces. While catalytic and many structural protein subunits of these complexes have been identified, the mechanisms and basic determinants of substrate recognition are still poorly understood.

The current study defined relatively simple single- and double-stranded determinants for association and gRNA-directed cleavage. To this end, we used an electrophoretic mobility shift assay to directly score the association of purified editing

*Reprinted with permission from "Determinants for association and guide-RNA directed endonuclease cleavage by purified RNA editing complexes from *Trypanosoma brucei*" by A. Hernandez, A. Panigrahi, C. Cifuentes-Rojas, A. Sacharidou, K. Stuart and J. Cruz-Reyes. 2008. *Journal of Molecular Biology*. 381:35-48. Copyright © 2008 Elsevier, Ltd.

complexes with RNA ligands, in parallel with U.V. photocross-linking and functional studies.

The cleaved strand required a minimal 5' overhang of 12-nt and a ~15-bp duplex with gRNA to direct the cleavage site. A second protruding element in either the cleaved or the guide strand was required unless longer duplexes were used. Importantly, the single-stranded RNA requirement for association can be upstream or downstream of the duplex, and the binding and cleavage activities of purified editing complexes could be uncoupled. The current observations together with our previous reports (91,95) show that association, cleavage and full-round editing by purified editing complexes have distinct determinants that increase in complexity as these editing stages progress.

Finally, we found that the endonuclease KREN1 in purified complexes photo-crosslinks with a targeted editing site. A model is proposed whereby one or more RNase III-type endonucleases in editing complexes mediate the initial binding and scrutiny of potential ligands, and subsequent catalytic selectivity triggers either insertion or deletion editing enzymes.

Introduction

The majority of primary mRNA transcripts in the single mitochondrion of kinetoplastids, including species of *Trypanosoma* and *Leishmania*, are plagued with frame-shifts and stop codons. Protein-encoding sequences are produced via an extraordinary maturation process involving specific insertion and deletion of uridylates

at often hundreds of editing sites (ESs) in a single transcript. This process is catalyzed by megadalton multi-subunit assemblies known as L-complexes, 20S editosomes, or editing complexes that contain between 16 and 20 known subunits and target ESs specified by the partial complementarity of pre-edited mRNA (pre-mRNA) and guide RNAs (gRNAs). For recent reviews, see (42,85).

RNA editing has been recreated *in vitro* at single model ESs in either natural-like (49,50) or completely artificial (95) substrates. Early mechanistic studies indicated that all steps of deletion and insertion editing were catalyzed by distinct enzymatic activities (48,51,73,74,82). More recently, it was shown that a deletion cycle involves the consecutive action of endonuclease KREN1, 3' exo-uridylylase KREX1 and/or KREX2, and ligase KREL1 (58,63,67,69,73). Similarly, an insertion cycle involves endonuclease REN2 or REN3, terminal uridylyl transferase KRET2, and preferentially, ligase KREL2 (57,59,65,67,73). Yet, KREL1 may be used in absence of KREL2 *in vitro* and *in vivo* (67,72,73,245). Potentially, KREN1 and KREX enzymes could also help proofread misedited insertion ESs bearing extra Us; i.e., mis-edited insertion sites could be targeted and repaired by deletion editing (48). Additional observations also suggest that deletion and insertion activities may occur at individual ESs *in vivo*. Namely, RNAi of KREN1 down-regulates editing of Cyb and COII pre-mRNAs *in vivo*, which only contain insertion ESs11. Also, RET2 was shown to add Us at deletion sites *in vitro* (97).

Pre-mRNA/gRNA hybrids are proposed to form two helical regions flanking an internal loop. The downstream (relative to the scissile bond) "anchor" duplex directs endonuclease cleavage immediately 5' to it, whereas the upstream duplex is thought to

tether the cleaved 5' piece during U-specific processing and re-ligation. The mechanisms of substrate recognition in assembled editing complexes are currently been addressed (for a recent review see (88)).

Previous studies in our laboratory using purified native complexes have shown that secondary structure rather than sequence-specific features are primarily required for full-round insertion editing (91,95). In a completely artificial 43-nt pre-mRNA/gRNA model substrate with single-helical turns flanking the central loop, simple features of this loop were manipulated to interconvert sites between insertion and deletion editing. Important insights on the specificity of substrate association with purified editing complexes were obtained in competition studies using parallel U.V. photo-crosslinking and full-round catalytic editing assays. Such studies, using a single photo-reactive 4-thioU and a ^{32}P atom at targeted ESs, showed a preferential association of complexes with deletion and insertion substrates, particularly with the most efficient model substrate currently available for full-round editing (A6 pre-mRNA/D33 gRNA hybrid) (51,95). The native complexes also exhibited a level of non-specific binding to unrelated transcripts. Interestingly, ribose 2'-H substitutions on the downstream helix and gRNA-side of the central loop significantly inhibited both pre-mRNA cleavage and photo-cross-linking activities at a targeted ES. Furthermore, a single 2'-H substitution adjoining the scissile bond obliterated the endonucleolytic activity but had no effect on photo-crosslinking, suggesting that the ribose 2'-hydroxyl at this position is relevant for catalysis not association of editing complexes (95).

One of the photo-crosslinking subunits in assembled editing complexes was proposed to be KREPA2 (MP63) (94), which as several other subunits, contains conserved domains that predict interaction with nucleic acids (54,55). Studies of purified recombinant proteins established that KREPA3 (MP42), KREPA4 (MP24) and KREPA6 (MP18) exhibit RNA-binding activity (207,229,246), but their precise function in assembled editing complexes remains to be determined. KREPA4 and KREPA6 exhibited preferential binding to poly (U) homopolymers, suggesting a role in the recognition of the natural 3'-poly (U) extension of gRNAs. These recombinant proteins also showed a general low-affinity binding for RNA.

While previous photo-crosslinking analyses provided insights on the specificity of the editing enzyme/substrate association, absence of crosslinking with certain mutant substrates could not be interpreted with certainty. Furthermore, whether purified editing complexes form transient or stable ribonucleoprotein complexes (RNPs) with cognate substrates is unknown. In the current study we used an electrophoretic mobility shift assay (EMSA) to directly examine, for the first time, RNPs formed by purified editing complexes. We applied EMSA, photo-crosslinking and endonuclease analyses to define substrate determinants for association and endonuclease cleavage, the first catalytic step of RNA editing. Both single-stranded (ssRNA) and double-stranded (dsRNA) RNA were required for these two stages of editing, but ssRNA for association can be satisfied in different ways, whether or not endonuclease cleavage activity is observed. Importantly, the determinants for association and cleavage can be uncoupled, and the

determinants for endonuclease cleavage are more complex than for association but less intricate than for full-round editing.

Finally, we compared preparations of native and affinity-purified editing complexes in association and catalytic assays, and established that one subunit that photo-crosslinks at a targeted ES is the essential endonuclease KREN1. The subunit KREPA2 (MP63) was also confirmed to photo-crosslink. A model is proposed whereby recognition of basic determinants including those defined here, leads to a preferential association of editing complexes with potential substrates. Such initial interactions may precede subsequent specialized contacts that trigger catalysis by either deletion or insertion editing.

Methods

Synthesis and Labeling of RNA

The ES1-radiolabeled A6 mRNA substrate was prepared by splint ligation as described (94). All other RNAs were synthesized *in vitro* by the Uhlenbeck single-stranded enzymatic method (215) and gel-purified.

For the preparation of 5'-end labeled substrates, gel-purified RNA was dephosphorylated by treatment by alkaline phosphatase at 37°C for one hour, followed by addition of SDS, EDTA and proteinase K to a final concentration of 1.5%, 5 mM, and 40 μg/mL, respectively and additional incubation at 50°C for 30 minutes. RNA was

purified by phenol/chloroform extraction and precipitated with ethanol. 5 pmols of dephosphorylated RNA were incubated at 37°C for 30 minutes with [γ-^{32}P] ATP (1:2 ratio of 5'-ends to ATP) and T4 polynucleotide kinase and gel-purified. For 3'-end labeling, 5 pmoles of gel-purified RNA were incubated at 4°C for 12 hours with an equimolar amount of [5'-^{32}P] Cytidine 3', 5'-Bis (Phosphate) and 15 units of T4 RNA ligase in RNA ligase buffer (50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 10 μg/mL BSA, 50 μM ATP, 10 mM DTT, 2U/μL anti-RNase (Ambion), 10% DMSO) and gel-purified.

Cloning, Cell Culture and Transfection

ORFs were amplified from *T. brucei* genomic DNA, kindly provided by Larry Simpson. The primers for KREN1 were designed as reported (61). For KREPB5, the primers were: forward CCC aagctt ATG AGA CGG GCT GTG GTA CTC CGT AC; and reverse CGC ggatcc CCG CCC TCC CAG TGC CAG CGC AAC TA (Hind III and Bam HI sites are in small case letters, respectively). The amplified products using Pfu DNA polymerase were treated with HindIII, BamHI and ligated to the pLEW79TAP expression vector, kindly provided by Achim Schnaufer. Constructs were linearized with NotI and used to transfect *T. brucei* strain 29.13 as described (247). Selection of transfectants was applied with 2.5 μg/mL phleomycin. KREN1 and KREPB5 expression was induced with 100 ng/mL and 1 μg/mL tetracycline, respectively, and confirmed by immunoblotting with the PAP reagent (Sigma).

Purification and protein composition determination of Editing Complexes

Chromatographic Purification of RNA Editing Complexes

Mitochondrial extracts were prepared from procyclic *T. brucei* strain TREU667 as described (89,216). Editing complexes were purified from mitochondrial extracts by consecutive anion exchange and DNA-affinity chromatography as described (89,203).

Tandem Affinity Purification of RNA Editing Complexes

Four liters of culture at a density of ~2.0×107 cells/mL were pelleted and lysed in 25 mL of 10 mM Tris-HCl, pH 8.0, 150 mM KCl, 0.1% NP-40, 1% Triton-X-100 and one tablet of EDTA-free complete protease inhibitors (Roche) for 30 minutes on ice. Lysis was confirmed by microscopy. Lysates were spun at 6000 X g for 15 minutes and the clarified extract purified by sequential IgG and Calmodulin affinity chromatography as described (248).

Mass Spectrometric Analysis of Native RNA Editing Complexes

Proteins in gel bands and complex mixtures were identified by LC-MS/MS analysis as described (53).

Photo-crosslinking, RNA Cleavage, Electrophoretic Mobility Shift, Competitions and Adenylylation assays

All assays are variations of the standard editing assay in our lab which consists of a mixture of a pre-annealed mixture of 10 fmols ^{32}P-labeled RNA and 1.25 pmoles unlabeled gRNA, completed to 20 μL with MRB [25 mM Tris-HCl, pH 8, 10 mM $Mg(Oac)_2$, 10 mM KCl, 1 mM EDTA, pH 8, 50 μg/mL hexokinase and 5% glycerol]

and, if applicable, competitor RNA at the indicated molar excess relative to the ^{32}P-labeled substrate. The mixture was pre-equilibrated for 10 minutes at 26°C and 2 μL of peak editing or TAP fraction was added. Prior to the assays, quantitative annealing of the RNA pairs tested was confirmed in native gels [as in (95)]. The sample was incubated at 26°C for 10 minutes then treated in an assay-specific manner. For cross-linking, samples were irradiated for 10 minutes under a 365 nm UV lamp, treated with RNase A and RNase T1 (50 μg/mL and 125 units/mL final concentrations respectively) at 37°C for 15 minutes, supplemented with SDS loading dye and loaded onto an SDS-polyacrylamide gel. For mRNA cleavage, purified editing complexes were pre-treated with 10 mM tetrapotassium pyrophosphate, pH 8, in MRB, for 5 minutes on ice to inhibit ligase activity (73); after incubation the mixture was deproteinized and RNAs were resolved on denaturing polyacrylamide gels. For electrophoretic mobility shift assays, the reaction mixture was loaded directly (no loading dye) onto a 1.5% agarose gel in 0.5 X TBE (45 mM Tris-borate and 1 mM EDTA) and run for 2 hours at ~5 V/cm at 4°C. Following electrophoresis, the agarose gel was dried under vacuum. EMSA with site-specific labeled transcripts were significantly more sensitive and reproducible than with end-labeled substrates, since the splint-ligation method used to generate the former (see above) exclusively incorporates phosphorylated fragments. Only the parental A6 substrate and Pair-27 were site-specific labeled using synthetic donor fragments [e.g., as in Fig. 1; (94)], although 5'-end labeled A6 parental and other constructs were also compared side-by-side in shift assays. Immunodepletions were carried out as described for the immunoprecipitation of RNA cross-linking proteins (94)

using a monoclonal antibody against KREPA2 immobilized on goat anti-mouse IgG resin (Dynal). Adenylylation assays were performed as described (217). All assays can be scaled-up linearly to enhance signal. The data were reproducible in at least two independent experiments. Each experiment included repeat assays, and those shown are representative. Data were visualized by phosphorimaging and/or autoradiography.

Results

Our previous RNA-protein photo-crosslinking studies showed that purified native editing complexes preferentially associate with a model A6 substrate for full-round editing (Fig. 25A) via recognition of secondary structure not sequence-specific features (94,95). However, absence of cross-linking due to certain substrate modifications or reaction conditions leaves uncertainties about the editing enzyme/substrate association.

To directly score substrate binding by editing complexes, we established an electrophoretic mobility shift assay (EMSA). A standard reaction mixture for full-round editing or photocrosslinking studies, using purified editing complexes and an ES1-labeled substrate (Fig. 25A) (94), was briefly incubated and loaded onto a native agarose gel. A fraction of radiolabeled substrate exhibited delayed electrophoretic mobility only in the presence of editing complexes (Fig. 25B). This shifted product comigrated with complexes that were radiolabeled by adenylylation of ligase subunits (Fig. 25C) (217) and was specifically immunodepleted by monoclonal antibodies to editing subunits (Fig.

25D, upper). As expected, adenylylatable editing ligases were enriched in the antibody-conjugated IgG beads but not in beads without antibodies (lower).

To further confirm that these ribonucleoprotein assemblies (RNPs) include editing complexes, we examined their substrate specificity using competition analysis as those performed in photo-crosslinking and full-round editing studies (94). Importantly, the competition profiles in photo-crosslinking (that we reported in (94)) and EMSA assays were equivalent, i.e., the homologous A6 competitor was strongly inhibitory at 5–10 fold excess whereas tRNA and Cyb were significantly less inhibitory at 25-fold excess (Figs. 25E and 25F, respectively; and data not shown). Moreover, a similar competition pattern was observed in assays of gRNA-directed endonuclease cleavage, the first enzymatic step of a full-round editing cycle (Fig. 25G).

Together, these data indicate that the EMSA directly scores the editing enzyme/substrate association and specificity of editing complexes. The data using EMSA also mirrors the observations in parallel studies of RNA-protein photo-crosslinking and editing enzymatic activities. Furthermore, all these activities of editing complexes can be examined using common substrates and reaction conditions. Based on these observations, we sought to define substrate determinants for association and guide-directed cleavage by editing complexes.

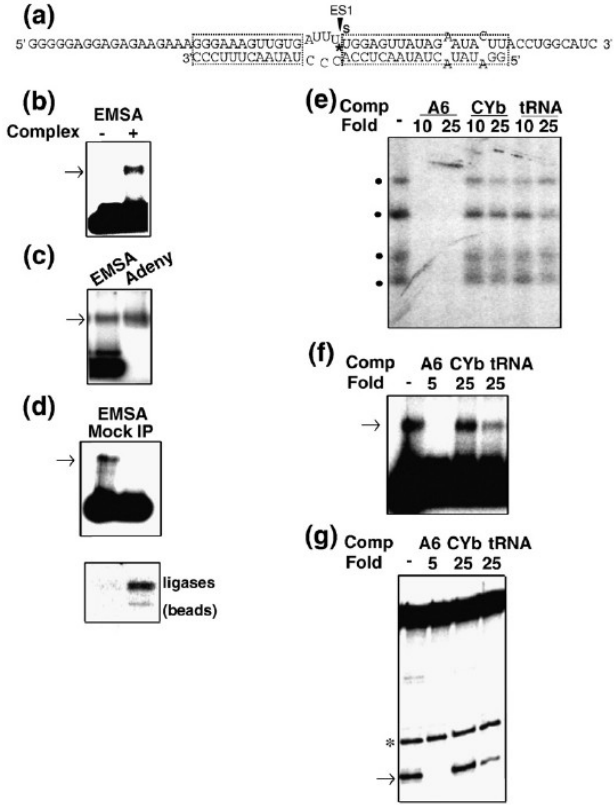

Figure 25. The association of purified editing complexes with substrates can be directly scored by EMSA, in parallel with UV photocrosslinking and pre-mRNA cleavage assays. (a) Scheme of an ES1-32P labeled (*) and thio-labeled (s) model A6 substrate for EMSA, UV photocrosslinking, and fullround U-deletion editing. (b) EMSA in a native agarose gel showing a shifted band (arrow) only in the presence of editing complexes. (c) Co-migration of the shifted substrate with editing complexes that were radiolabeled by auto-adenylylation (Adeny). (d) Specific depletion of the shifted product by immunoprecipitation (IP; upper) and recovery of selfadenylylatable ligase subunits in the beads (lower). A mock reaction was devoid of antibodies. (e) Preferential association of editing complexes with a substrate (A6) for fullround editing in competition studies using UV photocrosslinking (dots indicate four major crosslinks) or parallel assays of (f) EMSA and (g) endonuclease cleavage (arrow). A spurious cut (*) serves as loading control. In the EMSA, much of the substrate remained unbound. The fold excesses of unlabeled homologous A6 (5- or 10-fold) and heterologous Cyb and tRNA competitors are indicated. No competitor is (−).

We performed competition analyses, as in Figs. 25E–G, to examine the effects of unlabeled derivatives of the homologous (A6) competitor (diagramed in Fig. 26A). Our standard editing mixtures include gRNA at ~120-fold excess over radiolabeled A6 pre-mRNA to ensure quantitative annealing (95). In order to form competitor duplexes, the abundant free gRNA ("guide strand") in the standard mixture was allowed to pre-anneal with each pre-mRNA derivative ("substrate strand") added at a small, 5–10 fold, excess over radiolabeled pre-mRNA (Fig. 26B). All constructs in Fig. 2A used the same guide strand, and quantitative annealing was confirmed in native gels (95) (see methods).

Such analysis in binding and catalytic assays performed in parallel is illustrated in Fig. 27. In this example, both the homologous A6 pair and derivative Pair-1, whose guide strand is fully based paired (i.e., it forms a continuous 33-bp duplex), were strong competitors in photo-crosslinking, EMSA and cleavage assays (Figs. 27A–C, respectively). However, a second derivative that conserves the 33-bp duplex but lacks overhangs (Pair-2) was a poor competitor in all assays. These data suggest that editing complexes associate with Pair-1 but not Pair-2. Thus the presence or absence of the central loop region in the parental A6 construct does not significantly affect the binding efficiency of editing complexes, although ssRNA seems required for association.

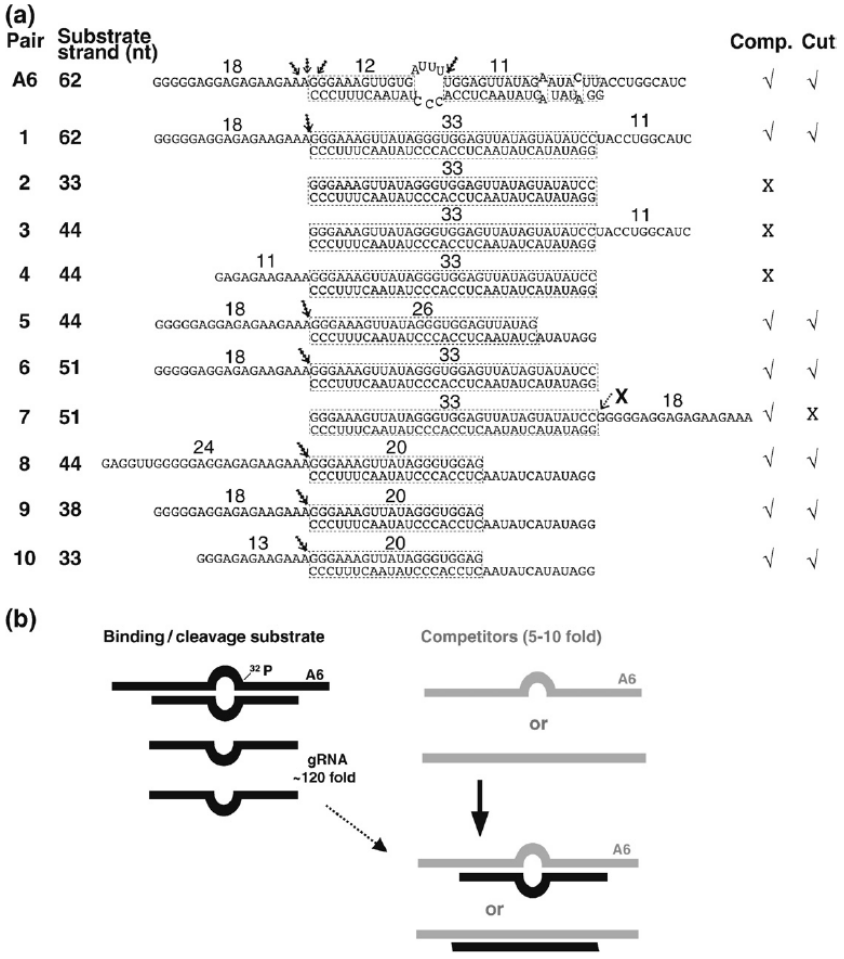

Figure 26. Constructs tested and scheme for competition assays. (a) Homologous A6 and derivative competitors (top "substrate" strand) paired with gRNA D33 (lower "guide" strand). The assigned number of each competitor RNA pair and size (nt) of the substrate strand are indicated, as well as the size of the predicted helix and overhangs. The demonstrated cleavage sites are noted with an arrow. Evident (✓) or weak-to-undetected (X) competition and cleavage activity for each construct are indicated at the right. Some pairs were not tested for cleavage activity. Pair 7 was tested for cleavage although a negative result was expected (see the text). (b) Cartoon of model RNA construct in standard functional (left) and modified competition (right) assays. 32P-labeled A6 pre-mRNA is usually annealed with complementary gRNA at ~120 fold excess in our standard editing assays. Unlabeled A6 substrate strand or variants (light strand) at 5- to 10-fold excess, over radiolabeled A6, anneal with free gRNA (both as dark strands), forming competitor pairs (light/dark hybrids).

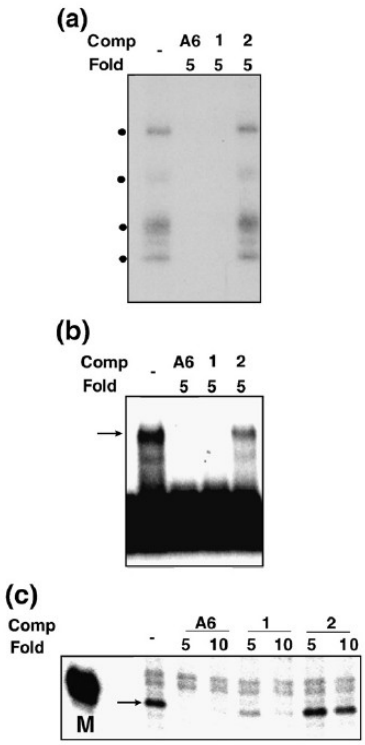

Figure 27. Parallel Competitions in UV Photocross-linking, EMSA and RNA Cleavage Assays. A6 and variant competitors (Comp) diagrammed in Fig. 26A were examined at the indicated fold-excess. Our cleavage assays typically include a size marker (M) such as the ^{32}P-kinased donor fragment used to prepare the parental A6 substrate (Fig. 25A) or control lanes with and without gRNA.

To dissect RNA requirements for association with editing complexes that distinguish Pair-1 from Pair-2, we designed competitors based on Pair-2 that contain upstream and/or substrate strand downstream overhangs of various lengths (Fig. 28; diagrammed in Fig. 26A). While 13-nt, 18-nt and 24-nt extensions favored association of editing complexes (Pairs 5–10), 11-nt extensions at either side of the duplex (Pairs 3–4) did not. Furthermore, constructs with shorter duplexes, 26-bp (Pair-5) and 20-bp long (Pairs 8–10), were also effective competitors. Most of these constructs used a 44-nt substrate strand, however, Pair-10 with a 33-nt substrate strand was also a significant competitor. Some competitions are more evident in cross-inking and EMSA than in cleavage studies (Figs. 28A–C). This difference may reflect different dynamics in the assays; that is, the former two score RNP complexes that either are present at the time of cross-inking or that withstand gel electrophoresis, respectively, whereas the latter scores accumulation of cleaved product over time, regardless of the relative stability of RNPs. Together, the competition studies in Fig. 25–Fig. 28 suggest that association with editing complexes requires recognition of a relatively simple structure bearing discrete ssRNA and dsRNA determinants.

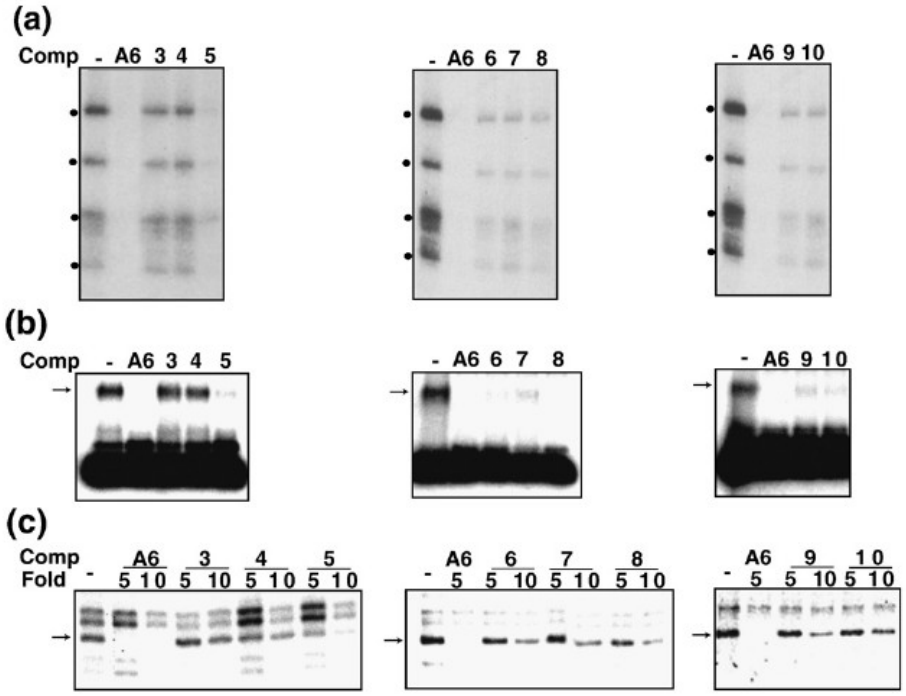

Figure 28. Parallel Competitions as in Fig. 27. (a–c) The homologous A6 and derived competitors are diagrammed in Fig. 2a.

Several constructs examined so far were effective competitors, indicating that are bound by editing complexes, but it was unclear whether they were also active in enzymatic assays. To directly address this, we tested these constructs for specific gRNA-directed cleavage by editing complexes (Fig. 29). Since the guide strand in these pairs fully complements the substrate strand we assayed for potential guide-directed cleavage at the phosphodiester bond just 5' of the duplex (50). We have reported that this particular bond is cleaved just 5' of the upstream duplex in the parental construct (Fig. 26A, top construct; and ahead in Fig. 29B) (51). Pairs 1, 5 and 6 generated a predicted 18-nt cleaved product (Fig. 29A) that corresponds to the 5'-end labeled overhang. This cleavage occurred only in presence of the guide strand. Furthermore, Pairs 8–10 which form a shorter 20-bp duplex were also cleaved with comparable efficiencies to the parental A6 construct (Fig. 29B). The expected 24-nt, 18-nt and 13-nt cleavage products, respectively, were gRNA-dependent. In the parental A6 construct, gRNA-directed cleavages occur 5' of both downstream (ES1) and upstream duplexes: the 5' end-labeled substrate strand accumulates a 31-nt product, as a result of consecutive cleavage and removal of 3Us by U-specific exonuclease activity at ES1 (63); also, multiple cuts 5' of the upstream duplex are observed probably due to misannealing of this helix. Spurious fragments of the substrate strand often accumulate due to breakage or RNase contamination that preferentially target Us in absence of guide strand, and are more evident with 5' labeled substrates.

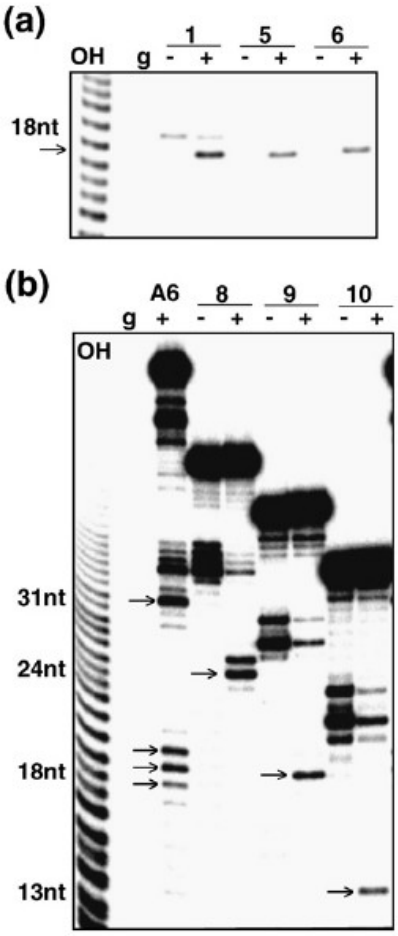

Figure 29. (a and b) Direct cleavage assays of 5′-End Labeled Substrate-strand Transcripts Paired with the Parental gRNA D33. The homologous A6 and derived competitors are diagrammed in Fig. 26a. Lanes with "+" and without "−" gRNA (g) are shown. Specific cleavage only occurs in the presence of gRNA (marked by an arrow). Spurious fragmentation of these transcripts occurs without gRNA but is inhibited by annealing of gRNA. Partial alkaline RNA hydrolysis "OH" was used as sizing ladder. Guide-directed cleavage of the A6 construct is directed by the downstream duplex (ES1) and by the upstream duplex. The latter occurs at three adjacent positions (~18-nt products) possibly due to alternative pairing. The short upstream duplex may be stabilized by coaxially stacking with the downstream duplex.

Among constructs found to associate with editing complexes, Pair-7 was not subject to guide-directed endonuclease cleavage as its substrate strand lacks a 5' overhang and its 3' ssRNA extension does not undergo cleavage (Fig. 26A; data not shown). The 18-nt protrusion of Pair-7 rescued the inactive Pairs 2–4 in crosslinking (Fig. 26A) and EMSA (data not shown) assays.

In summary, all efficient competitors in EMSA and photo-crossliking assays were also functional for endonuclease cleavage, except for Pair-7. While both association and endonuclease cleavage activities of editing complexes have ssRNA and dsRNA requirements, these could be present in a way that promotes association but not cleavage. Thus, association and catalysis by editing complexes can be uncoupled.

We decided to further analyze derivatives of Pair-10, the shortest construct tested that supported editing complex association and specific endonuclease cleavage activity. This symmetric construct with 13-nt overhangs flanking a 20-bp duplex was ideal to dissect determinants involved in selection of the substrate strand. That is, how are the substrate and guide strands distinguished in a duplex? We tested Pair-10 derivatives (Pairs 11–16) bearing progressively shortened 5'-overhangs in the guide strand (Figs. 30A and ahead in 30D). In these reduced structures, a 3-nt 5'-overhang in the guide strand promoted efficient cleavage of the substrate strand, but 1-nt and 2-nt extensions were strongly inhibitory (Pairs 15–16). Also, the latter constructs were not rescued by longer (18-nt) 5'-overhangs in the substrate strand (not shown). This suggests that 5' overhangs in the substrate and guide strands are not compensatory.

Analysis of constructs bearing shorter 5' extensions in the substrate strand (Pairs 17–20; Fig. 30B) showed that 12-nt are minimally required for endonuclease cleavage activity (Figs. 30D–E; and data not shown). Constructs with 11-nt 5'-overhangs in the substrate strand were inactive and not rescued by the presence of longer guide-strand overhangs (e.g., Pairs 19–20).

To determine whether constructs with duplexes shorter than 20-bp are functional we examined Pairs-21–26 (Figs. 30C and 30F). Efficient endonuclease cleavage was supported by Pair-21, which forms a 15-bp duplex, but progressive truncations of the guide-strand 5' overhang were increasingly inhibitory (Pairs 22–24). Pair-21 also showed that the substrate strand can be shorter that the guide strand, and that a ~27-nt substrate strand bearing a 12-nt 5'-overhang supports efficient endonuclease cleavage. In the above constructs the substrate-strand 5'-extension appears to be separately recognized, as inactivating truncations of this element were not compensated by a longer duplex or extended guide-strand 5'-ssRNA. In contrast, the guide-strand 5'-overhang could be replaced by using either an extended double-stranded terminus (e.g., Pair-6; Fig. 26A), or a 3'-overhang of the substrate-strand (Pair-25; Fig. 30C). The latter pair also showed that an 18-nt guide strand, largely annealed with the substrate strand, directs efficient endonuclease cleavage activity. Seiwert et al. reported that an 18-nt guide strand directs endonuclease cleavage of a complementary 73-nt A6 mRNA (50). Pair-25 and Pair-5, both of which generate the same cleaved product, were nearly as efficient as the parental A6 construct (Fig. 30G; see also Fig. 29A). Finally, we found that an 11-bp duplex in Pair-26 failed to direct detectable cleavage of the substrate strand (not shown).

Such 11-bp duplex seems relatively stable (−18.4 kcal/mol) and we confirmed efficient annealing with the substrate strand in native gels (95). Although this simple pair is not cleaved, it binds editing complexes in an EMSA (see the site-specific labeled Pair-27 in Fig 31A). Importantly, the ssRNA overhang was essential for binding, and substitution of a paired strand with DNA was inhibitory (Pair-28 and Pair-29, respectively). We examined additional constructs for association, whether or not they are cleaved, (Fig. 31B). In this case, we prepared derivatives of the thiolated parental A6 (diagrammed in Fig. 1) and tested their ability to photo-crosslink with editing complexes. For example, Pair-30 photo-crosslinks and is also cleaved (Fig. 31B; and data not shown). Other derivatives with an ssRNA overhang that crosslinked are not cleaved, whereas a blunt helix did not exhibit detectable crosslinking (Pairs 31–33, respectively). The parental A6 substrate generates more robust signals in association assays that most derivatives tested in our study.

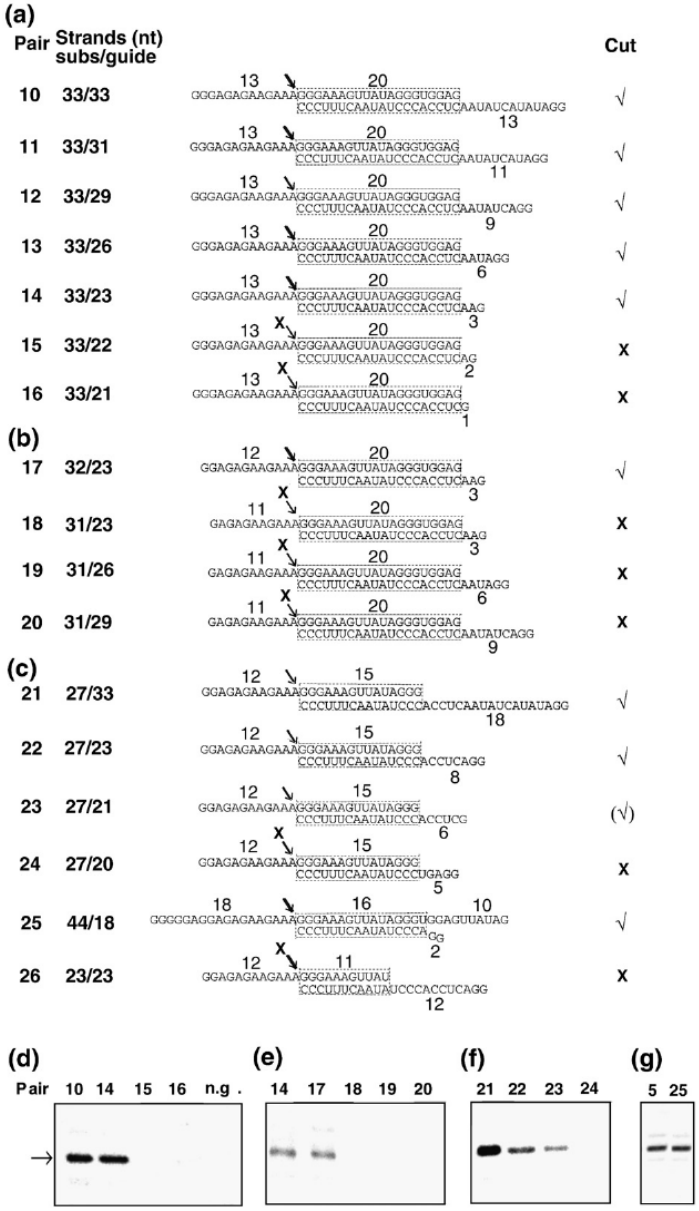

Figure 30. Diagram of minimized substrates for endonuclease cleavage by purified editing complexes. (a–c) A6 and derivative competitors (substrate strand) paired with parental gRNA.D33 or shorter versions (guide strand). The size of both strands in each pair is indicated. All other labels are as those in Fig. 26a. Detected (✓) or undetected (X) cleavage activity is indicated for each construct. Cleavage activity on Pair 23 was relatively weak. (d–g) Cleavage assays using 3′-end-labeled substrate strand derivatives.

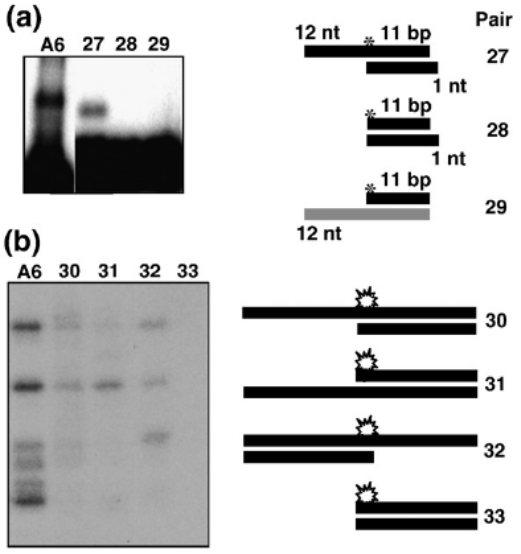

Figure 31. Additional RNA pairs that associate with purified editing complexes but are not cleaved. (a) EMSA; Pair 27 (derived from Pair 26) forms an RNP but is not cleaved. This RNP exhibits a faster electrophoretic mobility compared with the parental A6, but the reason for this is unclear. Duplexes without the 12-nt overhang or bearing a DNA strand failed to form an RNP (Pairs 28 and 29, respectively). (b) UV photocross-linking assays of the A6 parental construct in Fig. 1 and derivatives with or without an ssRNA overhang (Pairs 30–33). The site-specific ^{32}P label in (a) and the ^{32}P and thio labels in (b) are depicted by an asterisk and a star, respectively.

In summary, the construct series in Fig. 30 showed that purified editing complexes only cleave substrate strands bearing a minimal 5'-overhang of 12-nt. The minimal duplex directing specific cleavage was not determined to the nucleotide but it could be ~15-bp long, if not smaller. In addition to these two features, cleavage activity required the presence of (a) either a substrate 3'-overhang or a guide 5'-overhang when using a 15-bp duplex, or (b) a larger duplex without additional ssRNA. Fig. 31 confirmed that association and cleavage can be uncoupled although an ssRNA overhang is essential for both these two stages of editing. Importantly, association exhibits simpler determinants than cleavage.

It is feasible that some if not all determinants defined in the current study may be recognized by one or more RNA-binding subunits of editing complexes, including RNase III-type, OB-fold and zinc-finger domain-bearing subunits. At least three RNase III-type endonucleases identified in editing complexes are thought to catalyze pre-mRNA cleavage in insertion and deletion editing (57-59,69). However, the composition of the native editing complexes used here, including the presence of reported endonucleases, was unclear.

A mass spectrometric analysis of this protein preparation revealed nearly all reported subunits of affinity-purified ~20S editing complexes in *T. brucei* and *L. tarentolae* (58,59,69), in addition to subunits of the MRP complex which are thought to transiently associate with ~20S editing complexes via an RNA linker (Fig. 32A) (47). Three other proposed editing subunits, KREPA5, KREPA6 and KREH1, were not detected likely because they were either sub-stoichiometric, insufficiently ionized in our

preparation or absent. However, KREPA6 was recently reported to be essential (246) and was most likely undetected in our samples.

Since our previous photo-croslinking studies indicated that at least four subunits of purified ~20S native complexes make intimate contact with model editing sites (Fig. 32B) (94,95) we attempted the identification of a cross-linking subunit that migrates at about 100 kDa, where the endonuclease KREN1 was expected. To this end, we made a TAP-KREN1 construct and expressed it in *T. brucei* procyclic cells (see Materials and Methods section) based on a reported protocol that generated the same cell line (61). Tagged-editing complexes were purified through IgG and calmodulin-binding peptide (CBP) coupled resins and then examined by photo-crosslinking. We found that CBP-KREN1 complexes produced a shift of the ~100 kDa crosslink due to the mass added by the tag (~5kDa; Fig. 32C). These complexes also exhibited the crosslink by endogenous KREN1 and the other major crosslinks observed in native complexes. As far as we know this is the first evidence that at least two copies of KREN1 are present in editing complexes. Previous characterization of KREL1, KREN2 and KREN3 (KREPB2) affinity-purified complexes showed that endogenous and ectopic copies of these subunits were also present (47,61,67). Importantly, the shifted crosslink is specific of our tagged-KREN1 cell line, and not associated with the cell culture or protein purification conditions, as affinity-purified complexes using a different tagged subunit (TAP-KREPB5; i.e., MP44) exhibited the same cross-linking pattern of native complexes (Fig. 32C), as well as a similar silver staining pattern (Fig. 8D) and full-round insertion and deletion editing activity (not shown).

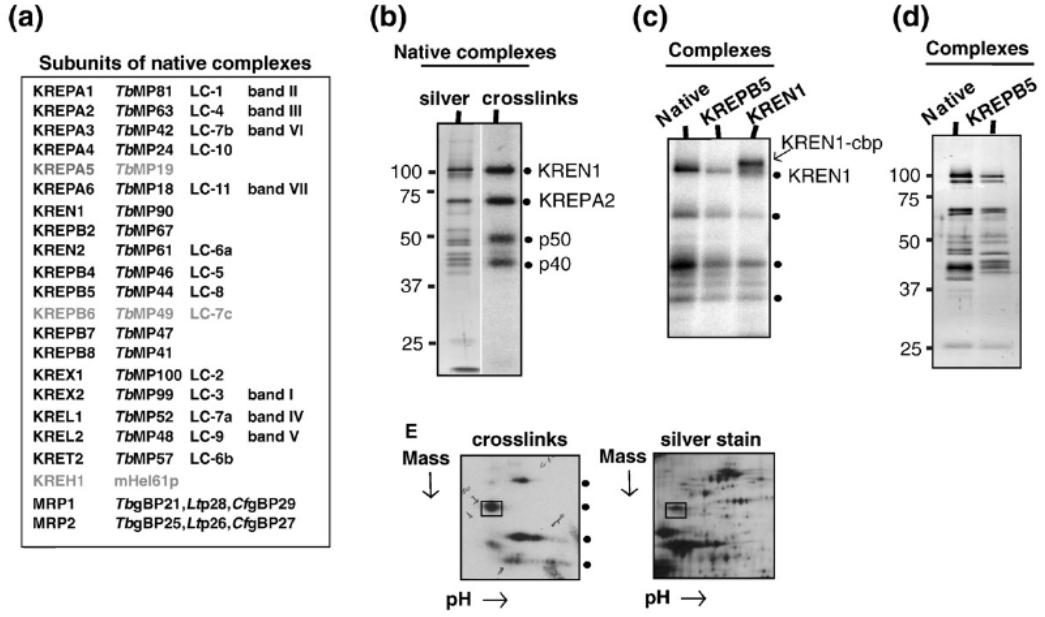

Figure 32. Composition of native editing complexes and identification of two photocross-linking subunits: RNase III-type endonuclease KREN1 and structural KREPA2 (MP63). (a) Listing of all subunits detected by mass spectrometry. Alternative nomenclature used in the literature is indicated. Three subunits were not detected (faded). (b) Native editing complexes stained with silver (lane 1) or exposed onto an X-ray film after UV photocrosslinking (lane 2). The cross-links (dots) by KREN1 and KREPA2 and two more subunits to be identified (p50 and p40) are indicated. (c) Cross-links by native (lane 1) or affinity-purified KREPB5 (MP44) (lane 2) and KREN1 (lane 3) complexes. Both CBP-tagged (upshift) and endogenous KREN1 are indicted. (d) Silver staining of native and affinity-purified KREPB5 complexes. This panel was prepared using complexes purified during this study (see Methods). Preliminary studies using aliquots from KREN1 and KREN2 complexes characterized in a previous study showed that only the former generate the 100-kDa crosslink (see the text). (e) Two-dimensional gel of partially purified complexes after photocrosslinking (left) or silver staining (right). Cross-linked KREPA2 (boxed) was excised from the gel and identified by mass spectrometry.

Consistent with the identification of KREN1 in the current study, our preliminary cross-linking analysis using aliquots of KREN1 and KREN2 complexes purified and characterized in another study (61) showed that the former but not the latter forms the 100-kDa crosslink (data not shown). The presence of these KREN proteins was mutually exclusive in the reported purified complexes (61).

Our previous 1D-analyses suggested that the crosslink at ~60 kDa was KREPA2 (94). We confirmed this identification by performing a 2D-gel analysis of a partially purified protein preparation exhibiting significant cross-linking activity by editing complexes (Fig. 32D, lower panel). The ~60-kDa crosslink was resolved in a discrete region of the gel, and mass spectrometric analysis of the excised region only contained KREPA2. The cross-linking subunits at about 50 and 40 kDa were more disperse and mass spectrometric analyses of these gel regions were unsuccessful. Thus, they remain to be identified.

Overall, the native editing complexes used in the current studies contain most subunits previously observed in purifications by other labs including the RNase III-type endonuclease KREN1, which we showed directly photo-croslinks with model editing sites. This subunit may be involved in the editing complex recognition of the substrate determinants defined here for association and endonuclease cleavage, but additional work is needed to explore this possibility.

Finally, we compared the substrate specificity of native editing complexes and KREPB5 affinity-purified complexes in parallel EMSA, photo-crosslinking and endonuclease cleavage assays (Fig. 33A–C). Native and tagged-KREPB5 editing

complexes exhibited similar substrate specificity, in presence of the homologous A6 (5-fold excess) and tRNA (25-fold), as positive control and relatively poor competitors, respectively. Thus, the approaches adopted in these studies should be useful in further comparisons of native and affinity-purified editing RNPs that exhibit different protein and functional composition.

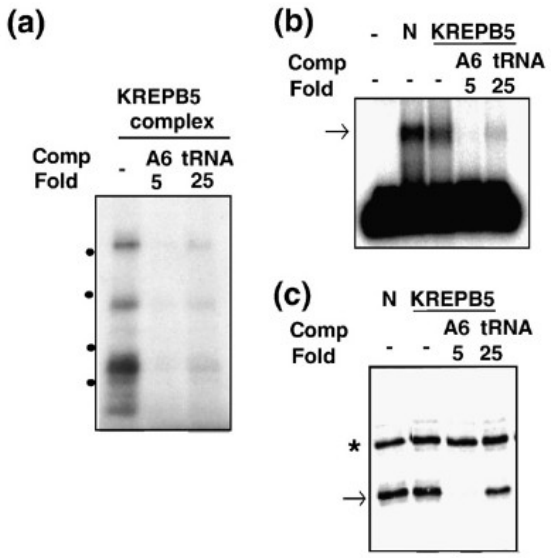

Figure 33. Association and endonuclease cleavage activity of affinity-purified editing complexes. Parallel (a) photocross-linking assay, (b) EMSA, and (c) cleavage assay. All labels are as those in Fig. 25. KREPB5-tagged complexes were directly compared with native "N" complexes.

Discussion

The goal of the current work was to define substrate requirements for association of purified editing complexes and gRNA-directed cleavage, the first catalytic step of an editing cycle. To this end, we used an EMSA, for the first time, in parallel with U.V. photo-crosslinking and gRNA-directed cleavage assays. Importantly, these assays were performed under comparable reaction conditions and the data obtained was complementary. The RNP assemblies detected by EMSA contained adenylylatable ligases and co-immunoprecipitated with known editing subunits (Figs. 25B–D), and their substrate specificity was conserved in the association and catalytic assays (Figs. 25E–G).

Our combined EMSA, photo-crosslinking and enzymatic studies defined ssRNA and dsRNA determinants for association and cleavage, summarized in Fig. 34. Three main combinations of ssRNA and dsRNA determinants supported endonuclease cleavage are represented by the following pairs: Pair-22 (27-nt substrate and 23nt guide strands) exhibits minimal 5' overhangs and ~15-bp duplex for cleavage. In this context, a 12-nt 5' overhang in the substrate strand was minimally required, whereas truncations of the 8-nt 5' overhang in the guide strand were gradually inhibitory. The size of one overhang did not compensate for the size of the other, and thus appear to involve separate recognitions. Pair-25, a long substrate-strand annealed to a minimal guide-strand of 18-nt (16-nt in a duplex) supports efficient cleavage. This confirms the observation by Seiwert et al., that an 18-nt guide strand directed endonuclease cleavage of a complementary 73-nt A6 mRNA 3 (50). Thus, a substrate 3'-overhang can substitute for a guide 5'-overhang. Pair-6, a long duplex overrides a requirement for

ssRNA rightward of the duplex. Thus neither these overhangs are essential but an ssRNA extension, abutting a short duplex, may suffice. In this type of construct, the size of the substrate 5'-overhang was also tested. 12-nt or more supported cleavage (e.g., Pair-6, and data not shown) but 11-nt was inactivating (i.e., Pair-4; data not shown). Additional pairs were bound but not cleaved by editing complexes, showing that these two aspects of editing can be uncoupled. Pair-27 is the simplest construct of this kind. Competition studies or straight association assays by crosslinking or EMSA showed that pairs bearing blunt-ended helices or insufficient ssRNA cannot associate with editing complexes. Pair-2 and Pair-28 reproducibly failed to form detectable RNPs and Pair-3 was significantly less effective than the parental A6 substrate (data not shown). Some constructs that bind but are not cleaved were examined by photo-croslinking or EMSA using 5'-end labeled rather than more sensitive site-specific labeled RNAs (Fig. 31B; and data not shown).

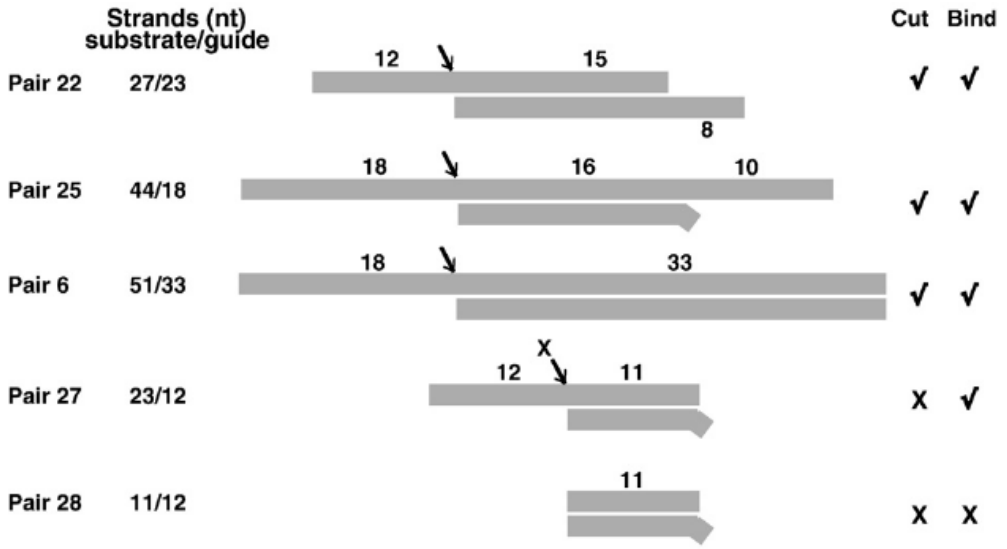

Figure 34. Summary of defined ssRNA and dsRNA determinants for endonuclease cleavage and association by purified editing complexes. Important variations were observed depending on the secondary structure context. Three main types of cleaved constructs are illustrated by Pair 22: It bears minimal substrate 5′ and guide 3′ overhangs. In this context, further shortening of either element was strongly inhibitory and not rescued by lengthening of the other. Pair 25: Its long substrate strand allowed reducing the guide strand to 18 nt. Thus, a substrate 3′ overhang can substitute for a guide 5′ overhang. Pair 6: Its long duplex can substitute for either substrate 3′ overhang or guide 5′ overhang. Thus, neither of these overhangs is essential but one may suffice in cleaved constructs. Importantly, association can occur without cleavage although it also requires an essential overhang either upstream or downstream of the helix. This is illustrated by Pair 27 and Pair 28. Detected (✓) or undetected (X) binding (bind) and cleavage (cut) are indicated.

Together, these constructs as well as others examined indicated that an appropriate combination of dsRNA and ssRNA determinants, rather than overall size of the bi-molecular structure, is required for both association and endonuclease cleavage by purified editing complexes. The ssRNA requirement (12-nt) 5' of the scissile bond and the dsRNA/ssRNA combinations 3' of it seem to involve separate recognitions. The smallest helix tested that directed endonuclease cleavage was 15-bp long (~1.5 helices) but shorter versions similar to Pair-26 may be feasible (Fig. 30C). Although the shortest functional guide strand tested was 18-nt long, a functional guide strand may be longer than the substrate strand (e.g., Pair-21).

Importantly, the requirements for association and for catalysis can be uncoupled. This was shown by Pair-7 (Fig. 26), Pair-27 and A6 thio-lated derivatives (Fig. 31) that bind editing complexes but are not cleaved. In Pair-7, the substrate-strand forms a 3'-overhang but not a 5'-overhang. Its substrate-strand 3' ssRNA stimulates association (compare with the inactive Pair-2) but, as expected, is not cleaved since editing endonucleases specifically target the phosphodiester bond immediately 5' of the guiding "anchor" duplex (48,50). On the other hand Pair-27 bears the critical 12-nt 5'-overhang but either insufficient duplex or overall length for cleavage. Furthermore, while a substrate 5' 12-nt overhang is minimally required for cleavage, whether all residues need be unpaired or some may partially complement apposing guide-strand residues was not determined. In full-round editing substrates, single-strandedness of residues near the downstream "anchor" duplex is strongly stimulatory. More distal residues can engage in formation of a proposed upstream "tether" duplex in deletion or insertion *in vitro*

(51,93). Furthermore, the presence and/or the nature of gRNA residues in the internal loop may stimulate full-round editing. Consistent with this idea, the lack or inappropriate number of such residues inhibited full-round deletion and insertion editing (51,93) (and unpublished data), and 2'-deoxy substitutions on the gRNA-side of the internal loop inhibited both photo-crosslinking and cleavage at the scissile bond (95).

Previously, our lab defined a minimal 43-nt pre-mRNA/gRNA hybrid for efficient full-round editing, which formed 10-bp helices flanking the ES. These nearby helices may be stabilized by coaxial stacking interactions, resembling a continuous helix. The smaller hybrid for endonuclease cleavage activity (including a ~27-nt substrate strand) and even simpler structure for binding imply that editing complexes require more extensive RNA contacts for the complete editing reaction, than for the intermediate cleavage step and the initial association step. Consistent with this concept, the artificially enhanced A6 parental substrate (51) is more efficient in all EMSA, photo-crosslinking and cleavage assays than most simpler derivatives tested here.

The fact that only one shifted product is reproducibly detected in the EMSA of the constructs examined suggests binding by a single editing complex, whether dimeric or of higher-order composition consistent with the co-purification of endogenous and ectopically expressed editing subunits, i.e., KREN1 in the current study (Fig. 32C) and KREN2, KREN3 (KREPB2) and KREL1 in previous studies (46,47,67). A mass spectrometric analysis revealed that the native complexes used in this study contain most known subunits of catalytic ~20S editing complexes, as expected from similar biochemical purifications (53). In addition, we found subunits of the MRP sub-complex

as it was reported in purified L-complexes (47,55), suggesting that at least some purified assemblies represent holoenzyme rather than core complexes. Several observations lead us to suggest that some if not all determinants defined in this study may be recognized by one or more RNase III-type proteins. Namely, (a) the shortest duplex tested that directed efficient endonuclease cleavage activity spanned ~1.5 turns. This is also the size of the smallest substrate identified that binds bacterial RNase III (149); (b) the critical role of 5' and 3' overhangs for cleavage at ssRNA-dsRNA junctions by the RNase III family member Drosha (249), and (c) the fact that KREN1 photo-crosslinks with a site for full-round editing (Fig. 32). This photo-crosslink was defined at a deletion site (Fig. 8C) but most likely also corresponds to a co-migrating crosslink at insertion sites (95). KREN1 endonuclease was proposed to specifically cleave deletion sites (58), however since association and cleavage are uncoupled we propose a model whereby KREN1 and related RNA-binding subunits may help scrutinize potential ligand determinants in the earliest checkpoint of RNA editing. Subsequent to the binding step, catalytic selectivity based on additional specific substrate recognitions may activate either the deletion or insertion enzymes, including proofreading of mis-edited insertion sites. A role of REN1 in an early checkpoint of ligand binding may explain why KREN1 down-regulation inhibits editing of Cyb and COII pre-mRNAs *in vivo*, which only have insertion sites. It is known that bacterial RNase III can undertake a modulatory role as a general dsRNA-binding protein regardless of its catalytic action (250). Importantly, the crosslinking activity of KREN1, KREPA2 (MP63) and at least two other major cross-linking subunits is conserved in both native and tap-tagged affinity-purified complexes. Such

conservation further suggests that the interactions are relevant, and independent of purification protocols and cell lines used. The conserved OB-fold and zinc fingers of KREPA2 may also be involved in recognition of single-stranded determinants defined here.

Finally, while the current study shows that RNPs formed by purified editing complexes can be directly visualized, it is currently unclear if the fraction of substrate that remains unbound in association assays reflects the concentration and/or affinity of either total complexes or functional complexes. Also, not all RNPs formed in solution may be stable enough to withstand the forces of gel electrophoresis. These and related questions will be addressed in separate studies.

CHAPTER V

REH2 RNA HELICASE IN KINETOPLASTID MITOCHONDRIA: RIBONUCLEOPROTEIN COMPLEXES AND ESSENTIAL MOTIFS FOR UNWINDING AND GUIDE RNA BINDING*

Summary

Regulation of gene expression in kinetoplastid mitochondria is largely post-transcriptional and involves the orchestration of polycistronic RNA processing, 3' terminal maturation, RNA editing, turnover and translation, however these processes remain poorly studied. Core editing complexes and their U-insertion/deletion activities are relatively well characterized and a battery of ancilliary factors has recently emerged.

This study characterized a novel DExH-box RNA helicase, termed here REH2 (RNA editing associated helicase 2), in unique ribonucleoprotein complexes that exhibit . unwinding and guide RNA binding activities, both of which required a double-stranded RNA-binding domain (dsRBD) and a functional helicase motif I of REH2.

REH2 complexes and recently identified related particles share a multi-protein core but are distinguished by several differential polypeptides. Finally, REH2 associates

*Reprinted with permission from "REH2 in kinetoplastid mitochondria: ribonucleoprotein complexes and essential motifs for unwinding and guide RNA binding" by A. Hernandez[&], B.R. Madina[&], K. Ro, J.A. Wohlschlegel, B. Willard, M.T. Kinter and J. Cruz-Reyes. 2010. *The Journal of Biological Chemistry.* 285 (2):1220-1228. Copyright © 2010 The American Society for Biochemistry and Molecular Biology, Inc. [&] Equal contribution

with ancilliary factors that control editing and RNA stability.

We propose that these putative transiently, via RNA, with editing complexes, mitochondrial ribosomes and severalhigherorder structures coordinate mitochondrial gene expression.

Author Contributions

Pages 154-187 are the original text and figures reported in reference #109. REH2 RNA helicase in kinetoplastid mitochondria: ribonucleoprotein complexes and essential motifs for unwinding and guide RNA (gRNA) binding. Hernandez A*, Madina BR*, Ro K, Wohlschlegel JA, Willard B, Kinter MT, Cruz-Reyes J. J Biol Chem. 2010, 285(2):1220-8. Hernandez and Madina share first-authorship (*). Author contributions are as follows:

Pg. 163, Fig. 35: Madina BR

Pg. 164, Fig. 36 (panels A,B,C,E) Hernandez A, (panel D) Madina BR.

Pg. 166, Fig. 37: Hernandez A, Panel C: Madina BR

Pg. 167, Fig. 38: Hernandez A

Pg. 169, Fig. 39: Madina BR

Pg. 171, Fig. 40: Hernandez A. The RNAi cell line in panels C-D was prepared by Madina BR

Pg. 173, Fig. 41: Hernandez A

Pg. 175, Fig. 42: Hernandez A

Pg. 187, Fig. 43: Cruz-Reyes J

Pg. 177, Table 5: Madina BR

Pg. 178 and 179, Tables 6 & 7: Madina BR

Pg. 179 and 181, Tables 8 & 9: Madina BR

(REH2 pull-dowsn form native complex preparations in tables 5-8, labeled "N": Hernandez A)

Pg. 181, Table 10: Madina BR (H2 TAP complex: Hernandez A); Ro K, Wohlschlegel JA, Willard B, Kinter MT performed the mass spectrometric studies in collaboration with Madina BR and Hernandez A.

Introduction

Unique gene expression mechanisms in kinetoplastid flagellates include U-insertion/deletion RNA editing by concerted cycles of cleavage, U-addition/removal and ligation that can create hundreds of amino acid codons in most mitochondrial mRNAs (42,85). The RNA editing core complex (RECC) contains 18-to-20 subunits (47,53,92,203) although a few subunits seem to exchange in substrate-specific variants of this complex (57). The RECC acronym was recently introduced by Larry Simpson et al. (56). Editing complexes recognize partial helices between pre-mRNA and complementary guide RNAs (gRNAs) initially stabilized by a short "anchor" duplex (50,92,227). Substrate determinants for duplex binding and nuclease-specificity (91,92,95) and substrate structure in solution (226,234,251) have been characterized.

Several accessory factors, mostly in multi-subunit arrays, have been proposed to modulate RNA editing during catalysis, substrate production or RNA turnover. The MRP complex has RNA annealing activity *in vitro* and may promote mRNA and gRNA pairing (210,211). Postranscriptional mRNA terminal 3'-poly A/U and gRNA 3'-poly U maturation is mediated by KPAP1 and RET1 complexes (128,252). MRB1, TbRGG1 and GRBC complexes proposed to contain between 14 and 24 proteins (termed here MRB-related complexes) share several components but their functional relationship remains unclear. Repression of a few common subunits inhibited RNA editing, and in some cases also decreased the level of total gRNA. GRBC1 and GRBC2 co-purified with RECC subunits (110,111,128,253-255). MERS1, MRP and RBP16 proteins were associated with mRNA stability (100,111). RBP16 also stimulated RNA insertion *in vitro* (103,104). DEAD-box mHel61 (also termed REH1) is the only predicted helicase known to impact RNA editing (114). Most of these proteins are likely to have additional roles outside editing. RNA helicases are common across species and typically multifunctional however only a few examples have been studied in mitochondria. This work characterized the protein and RNA interactions of a factor REH2 (Tb927.4.1500) that we initially found in native editing complexes of *T. brucei*. REH2 has a conserved dsRNA-binding (dsRBD) and DExH-helicase domains, and forms novel ribonucleoprotein complexes (RNPs) containing helicase activity, gRNA and a protein array that overlaps with MRB-related complexes. The integrity of REH2 RNPs and their helicase and gRNA-binding activities require the dsRBD. REH2 associates, via RNA, with RECC, a battery of accessory editing factors and mitochondrial ribosomes; thereby,

we propose that REH2 RNPs are integral components of RNA-linked supramolecular networks that orchestrate the expression of the mitochondrial genome.

Methods

TAP and RNAi Constructs, and Site-directed Mutagenesis of REH2

A TAP-REH2 construct was made by PCR amplification of the entire ORF from procyclic genomic DNA using a proofreading thermostable polymerase mix (AccuTaq-Sigma), and cloning into the Xho I and Bam HI sites of pLew79-ada-TAP. PCR-based site-directed mutagenesis was performed directly on this plasmid to alter the helicase motif I with oligonucleotides F-REH2-mI and R-REH2-mI, and delete the dsRBD with oligonucleotides FdsRBD-Δ and R-dsRBD-Δ. An RNAi construct was obtained by cloning an REH2 fragment of 1665 bp into p2T7-177 (256). All constructs were confirmed by DNA sequencing, linearized with Not I and transfected in procyclic 29-13 trypanosomes (247). REH2-TAP expression and RNAi were induced with tetracycline at 1 μg/ml.

Protein Purification and Analysis

Native editing complexes were purified by ion-exchange chromatography from mitochondrial extracts (92,203) and TAP-purifications were performed essentially as reported (248) with some modifications. Sedimentation fractions were obtained from

freshly made mitochondrial or whole-cell extracts in 10-30% glycerol gradients. While our protocols to prepare the extracts include Dnase I, a sample indicated in the text was subjected to an extra Dnase treatment (DNA-free kit, Ambion) prior to sedimentation. Catalase and thyroglobulin were used as ~10S and ~20S markers, and western blots of Tbmp45 (formerly termed REAP1) to determine the ~40S region (48). Affinity-purified REH2 antibodies were produced against the peptide CSHTPTTSAEAGGDS (Bethyl laboratories, Inc). IPs of endogenous and ectopic REH2 used antibody-conjugated protein A-dynabeads (Invitrogen). Ectopic REH2 was specifically immunopurified using anti-rabbit IgG dynabeads (Invitrogen). All washes were performed at 150 mM KCl. For mass spectrometry analyses the antibodies were cross-linked to the beads with 25 mM DMP (dimethylpimelimidate) in 0.2 M triethanolamine, pH 8.2.

Enzymatic Assays and Photo-crosslinking

The conditions and substrates to assay for full-round (88) and pre-cleaved (85) editing were as described. Photo-reactive substrates containing a single thio-U and ^{32}P at the editing site were prepared (88,94) and gRNA labeling was performed (227) as reported. RNA helicase assays used a dsRNA substrate consisting of the pre-mRNA fragment A6-tag annealed to the cognate gRNA gA6[14] (50). The dephosphorylated mRNA was 5'-end labeled with [γ-^{32}P]-ATP, and annealed with a 10-fold excess of gRNA in RNA folding buffer (25 mM Tris-HCl, pH 8.0, 250 mM KCl, 10 mM $Mg(Oac)_2$, 0.5 mM EDTA) by incubation at 95^{o}C for 10 min, followed by a gradual return to room temperature over the course of 2 hours. The annealed form was purified

by native gel electrophoresis in a 8% polyacrylamide gel in 1X TBE supplemented with 10 mM $Mg(Oac)_2$. The standard RNA helicase assay consisted of 50 cps (~10 fmols) of dsRNA in 25 mM Tris-HCl, 22 mM KCl, 10 mM $Mg(Oac)_2$, 0.5 mM EDTA, 3 mM DTT, 1 U/μL RNase inhibitor, 1 mM ATP, 50 ng/μL BSA, a 20-fold excess of an unlabeled trap ssRNA that complements ~33 bp of gRNA and 10 μL of beads in a final volume of 20 μL. Reactions were incubated for 30-60 min at 26^{o}C with constant flicking to mix the beads. This was followed by addition of 4 μL 6X stop solution: 0.12 % Xylene cyanol, 0.12% bromophenol blue, 3% SDS, 125 ng/μL proteinase K, 17 % glycerol and incubation at room temperature for 10 minutes. Samples were then loaded onto a 8% polyacrylamide gel, 1 X TBE supplemented with 10 mM $Mg(OAc)_2$. Protein RNA photo-crosslinking was performed as described (94) except that it was scaled-up 10-fold. Denaturation of complexes was accomplished by the addition of SDS to 1% final, and sequential incubations at 95°C for 10 min and at 70°C for 30-60 min. After allowing the sample to reach room temperature, triton-X-100 was added to 5%, and incubated 10 min at room temperature. Samples were then passed through a gel filtration spin column (Bio-spin 6, Bio-Rad 732 6221) according to the manufacturer's instructions and immunoprecipitated as above.

RNase Treatment of TAP Purifications and Immuno-precipitations

Samples coupled to Dynabeads where treated with the following nucleases as indicated in the text at the given final concentration: RNase A (0.1u/μL), T1 (0.125U/μL), V1 (0.001/μL) and micrococcal nuclease (0.03 U/μL) for 60 min in ice.

Mass Spectrometry

TCA precipitates were resuspended in digestion buffer (100 mM Tris-HCl, pH 8.5, 8M urea) and digested by the sequential addition of lys-C and trypsin proteases as described (257). Digested samples were fractionated using a 5- step online separation method during which peptides were eluted directly into a LTQ-Orbitrap mass spectrometer (Thermo Fisher) in which tandem mass spectra were collected (258-260). SEQUEST and DTASelect algorithms were used to identify peptides sequences from tandem mass spectra (261,262). Proteins were considered present in a sample if at least two peptides were identified per protein using a peptide level false positive rate of 5% as determined using a decoy database strategy (263).

Results

A DExH-box Helicase Associates with RNA Editing Complexes

In a mass spectrometric analysis of native editing complexes purified from *T. brucei* mitochondria we detected multiple unique peptides of most RECC subunits and the accessory MRP factors (92). However, we also found a single peptide for a 241-kDa protein, termed REH2, with highly conserved DExH-helicase domains, a double-stranded RNA binding domain (dsRBD) and an N-terminal mitochondrial import sequence (Fig. 35A). Western blot analyses clearly detected a ca. 250 kDa protein in native but not in tandem affinity-purified (TAP) editing complexes (Fig. 35B), although a weak signal was apparent in TAP-REL1 complexes (see the middle panel). This suggested a transient interaction of REH2 with RECC that is disrupted during high-stringency affinity purifications. Silver-staining of native editing complexes did not evidently detect components near 250 kDa suggesting that REH2 may be sub-stoichiometric relative to RECC subunits (Fig. 35C).

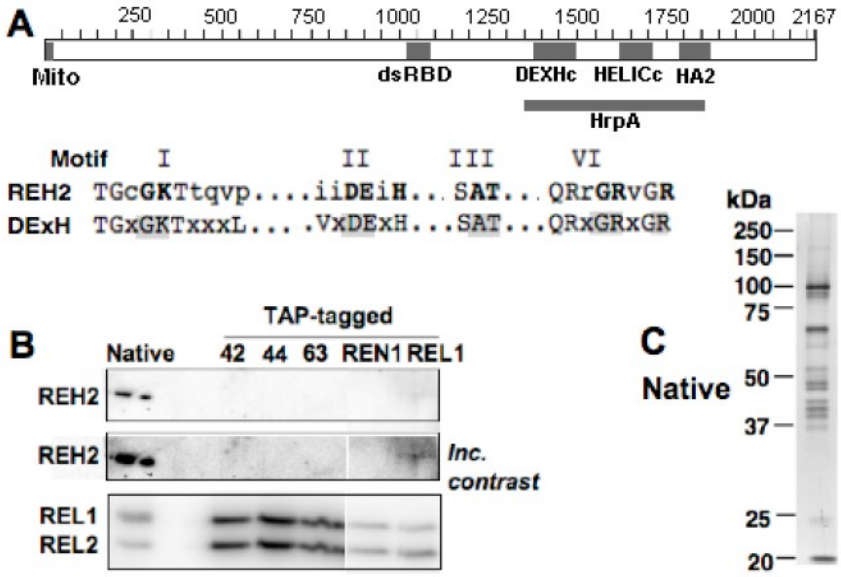

Figure 35. REH2 gene organization and co-purification with native editing complexes. (A) *T. brucei* REH2 has 2167 amino acids including a conserved mitochondrial import signal (Mito), double-stranded RNA binding domain (dsRBD) and domains typically associated with DExH-box helicases (drawn at scale): DexHc, HELICc, HA2 and HrpA. The first domain, which defines this protein family, contains six motifs including the motif II signature amino acids DExH where "x (205)" is any residue. Four such motifs are evident in REH2 (I, II, III and VI) and shown with the most conserved residues in gray (56); (B) REH2 western blot of native and TAP-purified editing complexes tagged at various subunits. The middle panel is shown at increased contrast. Adenylytated ligase subunits REL1 and REL2 shown as loading control; (C) Silver-staining of purified native editing complexes.

To confirm the physical interaction of REH2 with editing subunits, we expressed in procyclic trypanosomes the complete 6504-nt REH2 gene with a C-terminal TAP-tag [calmodulin binding peptide (CBP)-tobacco etch virus (TEV)-protein A (ProtA)] (248) under the control of a tetracycline-inducible promoter (Fig. 36A). REH2 mRNA and protein increased 5-fold and 2-to-3 fold, respectively, at day 3 of induction (Figs. 36B-C) and a slight reduction in cell growth was seen at day 5 (Fig. 36D). Consistent with its predicted mitochondrial localization, REH2 was enriched in a mitochondrial lysate (Fig. 36E) (205).

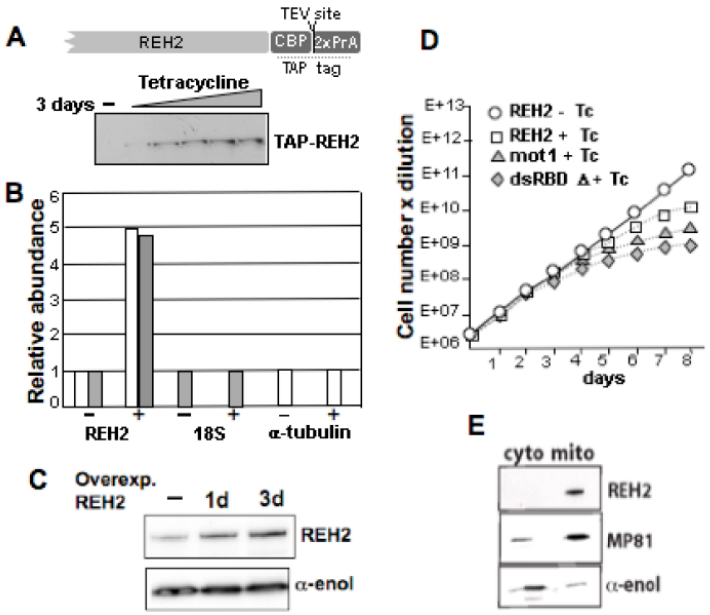

Figure 36. REH2 expression in procyclic *T. brucei*. (A) C-terminal TAP tag, and anti-CBP western of total cell extracts with or without 3 days of induction with increasing tetracycline (Tc) concentrations; (B) Quantitative RT-PCR of REH2 normalized to 18S rRNA and α-tubulin +/- Tc; (C) REH2 western of induced empty-TAP vector (-) and TAP-REH2 cells for 1 and 3 days. Cytosolic α-enolase was used as control. (D) Growth curves of procyclic trypanosome cell lines expressing TAP constructs of wild-type REH2, dsRBDΔ and motif I mutants. (E) Western blot showing REH2 enrichment in a mitochondrial extract. Mitochondrial MP81 and cytosolic α-enolase.

REH2 antibodies reacted well with purified native editing complexes and mitochondrial lysates but weakly with TAP pull-downs due to a partial occlusion of the tag and consequent low purification efficiency from mitochondrial lysates. A small amount of RECC was detected in both TEV (T) (Fig. 3A) and concentrated EGTA (ET) (Fig. 37B) eluates but not in mock pull-downs from empty-vector control cells (last lane in each panel). Interestingly, this low level of RECC in the eluates was resistant to a pre-treatment with Rnases A, T1 and micrococal nuclease (MN).

In a converse approach, REH2 was detected in mitochondrial extract pulldowns of several RECC subunits (Fig. 37C). Also, REH2 antibodies immunoprecipitated a small fraction of pre-isolated native editing complexes after an extensive RNase-MN treatment including the dsRNA-specific RNase V1. However, most RECC remained in the unbound fraction (Fig. 37D), suggesting that relatively few purified complexes were stably bound to REH2.

Besides the above examination of affinity-purified eluates and isolated native editing complexes, we further analyzed the REH2/RECC association in mitochondrial lysates. Importantly, while RECC subunits were present in nuclease-treated REH2 pulldowns, most RECC was released by the treatment (Fig. 37E). This suggests that the transient association observed is largely mediated by RNA. Also, in line with transient contacts, REH2 purifications exhibited some pre-cleaved but not full-round editing activity, which is less sensitive due to a limiting cleavage step (Fig 38).

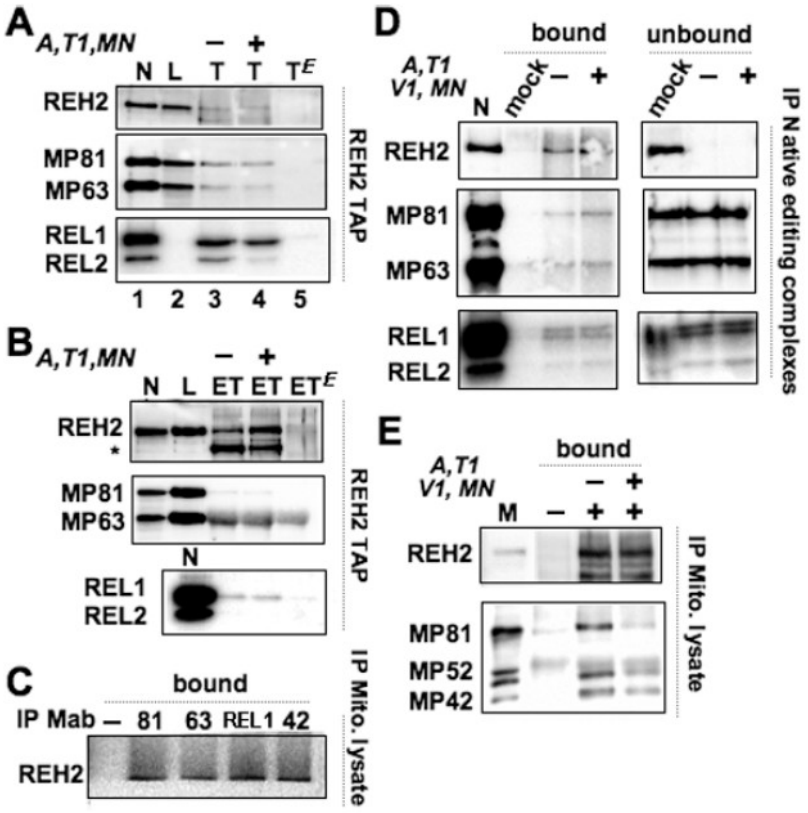

Figure 37. REH2 co-purification with RECC is largely sensitive to extensive nuclease treatments. REH2 TAP purification and detection of editing subunits before or after (+/-) treatments with Rnases A, T1 (and V1 as indicated) and micrococcal nuclease (MN) in (A) TEV (T) and (B) EGTA (ET) eluates from IgG-Sepharose and calmodulin-agarose, respectively. Control lanes include native editing complex (N), diluted whole-cell lysate (L) and empty-vector controls (TE or ETE; last lane). ET eluates were concentrated by acetone precipitation. Immunoblots of REH2 and editing subunits and auto-adenylylation of REL ligases. ATPases in whole-cell lysates often inhibit the latter activity (e.g., panel A, lane 2). MP63 may co-migrate with antibody cross-reacting BSA added as precipitation carrier (see the ETE control lane). REH2 fragmentation (*) is observed in the eluates. (C) Mitochondrial lysate IP pulldowns with antibodies to RECC subunits (MP81, MP63, REL1, and MP42) or a mock reaction. REH2 pulldowns from (D) purified native editing complexes or (E) mitochondrial lysate (M) and mock reaction.. An unaccounted band near MP52 is visible in the lysate lane.

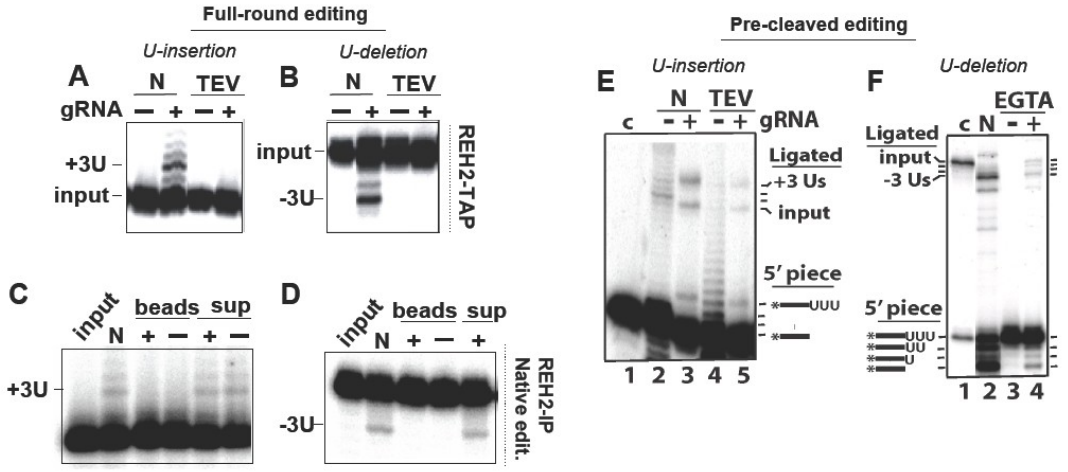

Figure 38. Pre-cleaved editing but not full-round editing activity was detected in the REH2 purifications. (A-D) Full-round U-insertion and U-deletion assays with accurately edited -3U or +3U products, respectively. (A and B) TEV eluates of a REH2 TAP-purification with or without gRNA; (C and D) REH2 IP pulldowns of purified native editing complexes. Positive control with native editing complexes (N). Beads or supernatants from pulldowns with REH2 antibodies (IP) or pre-immune serum (mock). (E and F) Pre-cleaved U insertion or U deletion assays showing: the 5' piece (bottom half of the gel) before editing or after addition or deletion of 3Us, respectively; Ligated RNA (top half of the gel) before or after U remodeling. 5' pieces are radiolabeled (*) and depicted as bars. U-insertion activity was detected only in the TEV eluate.

Heterodisperse REH2-associated Complexes are Stabilized by RNA and the dsRBD

Sedimentation analyses of mitochondrial lysates showed significant heterodispersion of REH2 with a broad peak at ~20-30S. Western blots of these fractions with REH2 antibodies and the PAP reagent (to score ectopic REH2 only) showed that the tag does not affect the sedimentation of REH2 (Fig. 39A, top and middle panels). The RECC subunit MP63 was also dispersed but in contrast to REH2 it was not detected in light fractions (Fig. 39A, bottom panel). Interestingly, REH2 at ~20S and >40S co-immunopurified with RECC subunits (see ahead Fig. 39D).

A pretreatment of the mitochondrial lysate with Rnases/MN disrupted most high S-value REH2 complexes generating a discrete peak at ~15S, whereas a significantly sharpened peak of editing complexes remained at ~20S (Fig. 39B). As described above, RNase treatment eliminated most REH2 association with RECC (Fig. 37E).

To examine the relevance of the conserved dsRBD we expressed a construct with a deletion of the entire motif (dsRBDΔ). Notably, this construct severely compromised cell growth (Fig. 36D) and reduced the sedimentation of both endogenous and ectopic REH2 to an extent comparable to the RNase treatment (Fig. 39C). Finally, we established that DNA does not largely contribute to the observed broad sedimentation of REH2 in a sample treated with Dnase (Procedures section; data not shown). Thus, REH2 forms heterodisperse particles that include editing complexes and are stabilized by RNA and the dsRBD, as well as relatively low-density particles that resist RNase.

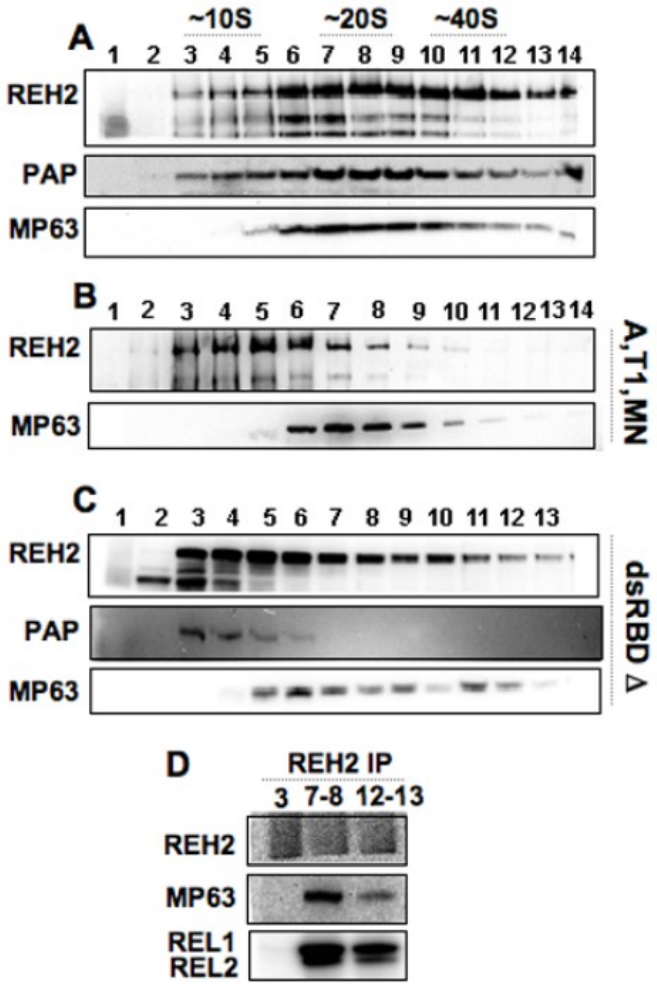

Figure 39. Heterodispersion of REH2 complexes requires RNA and the dsRBD. Glycerol gradients of mitochondrial extracts in western blots of endogenous plus ectopic REH2 (REH2) or ectopic REH2 (PAP reaction) and MP63 as a RECC marker. Mitochondrial lysates (A) untreated or (B) treated with Rnases A, T1 and MN, or (C) whole lysates of REH2 dsRBD-Δ cells. (D) REH2-antibody pulldowns of gradient fractions 3 (~10S), 7-8 (~20S), and 12-13 (>40S).

REH2 Co-purifies with gRNA in a Manner Dependent on the dsRBD and Helicase Motif I

Weng and collaborators recently reported that GRBC complexes, which co-purified with REH2, bind gRNA (111). We determined whether REH2 immuno-purified complexes associate with gRNA, and further examined the importance of conserved domains of this protein. To this end, we analyzed IgG-dynabeads pulldowns of ectopically expressed REH2 wild-type and mutants dsRBDΔ or motif I (GK-to-AQ) (Fig. 35A). The motif I mutated residues have been associated with ATP binding and hydrolysis in other DExH proteins (187).

While a significant amount of total gRNA co-purified with wild-type REH2, little if any was associated with either mutant (Figs. 40A-B). It is of interest, however, that relatively large RNA species accumulated in both mutants. Furthermore, as shown by Hashimi et al. (255), RNAi down-regulation of REH2 decreased the steady-state levels of gRNA (Figs. 40C-D). Importantly, REH2 pull-downs from wild-type cells contained gRNA, demonstrating that endogenous REH2 RNPs bind gRNA, and that this association is not an artifact of overexpression (Figs. 40E-F). Thus, gRNA-binding by REH2 RNPs requires the dsRBD and wild-type motif I of REH2.

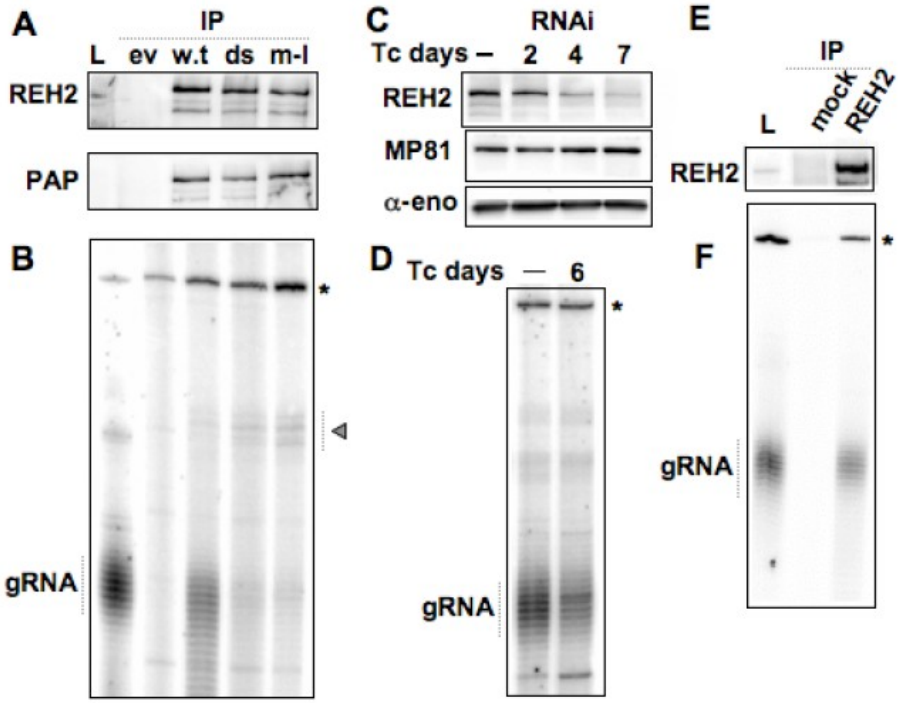

Figure 40. REH2 co-purification with gRNA requires the dsRBD and wild-type helicase motif I. (A) IgG-dynabead pulldowns (IP) of whole cell extracts expressing TAP-REH2: wild-type (w.t.) or mutants dsRBDΔ (ds) and motif I (m-I). Western blots with REH2 antibodies or the PAP reagent of pulldowns normalized for the amount of REH2; (B) gRNA-labeling assays with guanylyltransferase of the pulldowns in A. Whole-cell extracts from wild-type cells (L) or a REH2 pulldown from "empty-vector" cells (ev) were used as controls. A few RNA species accumulated in the REH2 mutants (arrowhead). A typical artifact in guanylyltransferase assays (*) serves as loading control; (C) Western blots of REH2 RNAi cells before or after the indicated days of Tc-induction. Mitochondrial MP81 and cytosolic α-enolase controls; (D) gRNA-labeling assays in REH2 RNAi cells before or after 6 days Tc-induction. (E) Western blot of a REH2 pulldown from 29.13 cells lacking the ectopic REH2 construct. A mock reaction with pre-immune serum, and whole lysate (L) are controls. (F) gRNA labeling of the pulldowns in E.: wild-type (w.t.) or mutants dsRBDΔ (ds) and motif I (m-I). Western blots with REH2 antibodies or the PAP reagent of pulldowns.

REH2 is Associated with RNA Helicase Activity and Appears to Photo-crosslink with RNA

We examined the REH2 IP pull-down from mitochondrial extracts for possible RNA helicase activity and RNA photo-crosslinking. Notably, a model A6 pre-edited mRNA, pre-annealed with cognate gRNA, was efficiently unwound by the REH2 pull-down in a reaction requiring ATP hydrolysis at its β–γ phosphodiester linkage (Fig. 41A). Although the dsRNA substrate in these assays was gel-isolated, some dissociation is visible in input and mock lanes in absence of REH2 complexes. A continuous duplex with 3' overhangs was also unwound (Fig. 41B) but not an identical helix with 5' or no overhangs (not shown), consistent with the substrate specificity of the vast majority of SF2 RNA helicases with the exception of DEAD-box proteins (183). Importantly, REH2 pull-downs from the dsRBDΔ and motif I cell lines had no detectable unwinding activity (Fig. 41C).

We analyzed the sedimentation distribution of this helicase activity relative to REH2 and the DEAD-box helicase REH1. Interestingly, most helicase activity sedimented in fractions containing REH2, but away from REH1 which localizes at the top of the gradient in fractions 1-3 (Fig. 41D-F) (114). An REH2 pull-down of fraction 8 exhibited unwinding activity (data not shown). Together with our above studies of REH2 mutants, this data suggest that REH2 is linked with the observed RNA helicase activity.

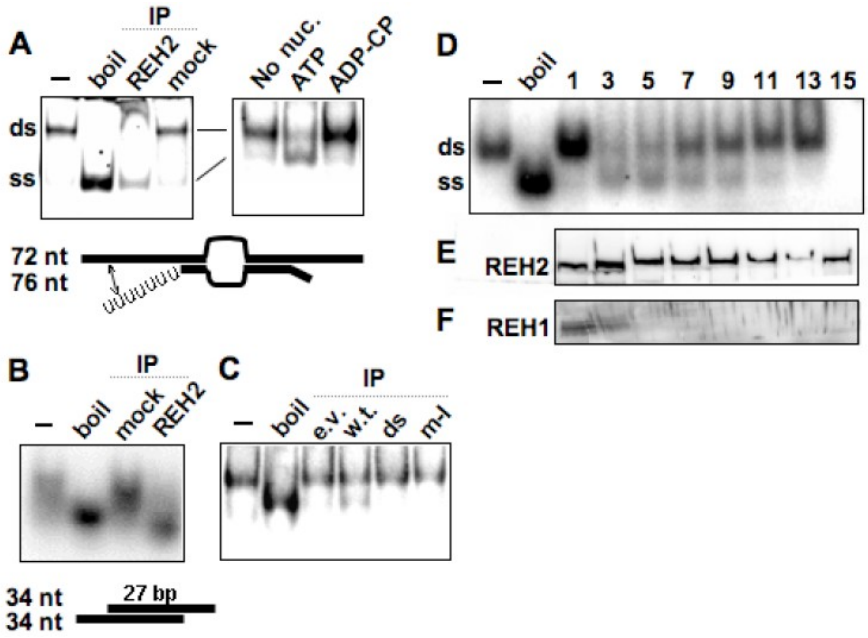

Figure 41. REH2 is associated with helicase activity. RNA helicase assays-REH2 antibody pulldowns from mitochondrial extract supplemented with gel-isolated dsRNA substrates. (A) A6 pre-mRNA and cognate gRNA with 1 mM ATP, ADP-CP or no nucleotide. The pre-mRNA was 5'-end radiolabeled. The gRNA 3'-oligo U hybrid is presumably unstable (arrow heads); or (B) a short duplex with symmetrical 3' overhangs and 1 mM ATP. (C) IgG-dynabead pulldowns of TAP-REH2 wild-type and mutants (see Fig 5 legend). RNA input before (-) and after boiling, and mock pre-immune serum pull-downs from empty-vector cell extracts are controls. (D-F) Sedimentation fractions of whole-cell lysate examined for helicase activity and western blots of REH2 and REH1.

To examine the possibility that REH2 may directly bind RNA we cross-linked a protein fraction enriched in RECC (203) with the pre-mRNA/gRNA substrate described above but substituted with a photo-reactive thio-U and ^{32}P at the editing site (94). This protein fraction produced a crosslink at ~250 kDa (Fig. 42A, lane 1) that was enriched in a pulldown with REH2 but not REL1 antibodies (lanes 2 and 3). Importantly, REH2 pulldowns of crosslinked reactions that were treated with SDS and increasing temperature to dissociate the RNPs further enriched the ~250 kDa crosslink (lanes 4-5), suggesting that the reacting protein is REH2. As a proof-of-concept for the above denaturation protocol (Fig. 42B), we isolated the RNA photocross-linked RECC subunits REL1 (lane 2) and reported MP63 (lane 3) (92,94) using specific antibodies against these proteins.

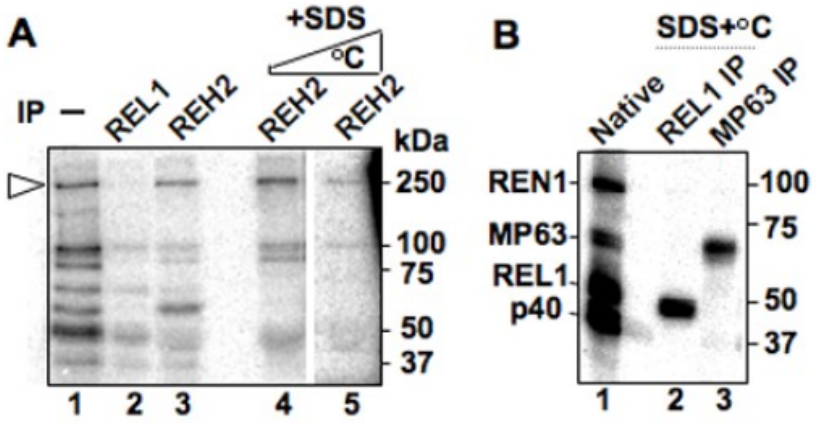

Figure 42. A ~250 kDa protein in REH2 pulldowns photo-crosslinks with RNA. (A) Crosslinking of an A6 pre-mRNA/gRNA pair bearing a single photo-reactive thio-U at the editing site. Mitochondrial Q-sepharose fraction enriched with RECC (lane 1) or IP pulldowns of this material by antibodies to REL1 or REH2 (lanes 2 and 3, respectively). Lanes 4-5 are repeats of lane 3 but after a treatment with 0.1 % SDS at 70°C and 90°C, respectively, that enriches a crosslink at ~250 kDa (arrowhead). The REL1 pulldown shows at least four reported cross-linking RECC subunits (92,94) (see panel B). (B) Proof of concept of the denaturation protocol used in panel A, showing 175mmune-purification of crosslinked RECC subunits. Crosslinked subunits of purified RECC (lane 1), and specific RECC subunits (lanes 2 and 3) after SDS dissociation at 90°C, and 175mmune-purification with the indicated antibodies. The efficiency of this procedure depends on the antibody's affinity, polyclonal for REH2 and monoclonal for RECC subunits.

Additional studies are needed to confirm that the ~250 kDa cross-link represents direct binding by REH2, but these data suggest that REH2 contacts the RNA duplex near the photo-reactive moiety in the model editing site (≤4Å) (264).

At Least Seven RNase/MN-resistant REH2 Associated Proteins are Also Present in MRB1, TbRGG1 and GRBC Complexes

We performed mass spectrometric studies of REH2 IP pull-downs of both purified native editing complexes and 20-30S sedimentation fractions (num. 7-10 in Fig. 39A), the latter before or after RNase/MN treatment. A large number of REH2-unique peptides were found in all pull-downs reflecting the relatively large size of this protein. The 20-30S fractions contained twenty-two proteins (besides REH2) previously found in one or more of the reported MRB1, TbRGG1 and GRBC complexes (110-112). These complexes significantly overlap but also exhibit important compositional differences (Tables 5 shows RNase-resistant REH2 interactions).

Table 5. RNase-resistant REH2-associated Proteins. (A) RNase-resistant REH2 associated proteins that were found in MRB1, TbRGG1 or GRBC1 complexes

A

Protein ID *T. brucei*	REH2 IP N	REH2 IP 20-30S fractions	REH2 IP 20-30S fractions *RNase*	mAb43 IP or TAP-tagging MRB1	mAb43 IP or TAP-tagging tbRGG1	mAb43 IP or TAP-tagging GRBC1/2	mAb43 IP or TAP-tagging GRBC1/2 *RNase*	Name/motif
Tb927.4.1500 &	25	36	25	+	+	+	+	**REH2**/DExH/dxRBD
Tb927.2.3800	1	2	3	+	+	+	+	GRBC2
Tb927.5.3010	1	7	2	+	+	+	+	Ribosomal_S2
Tb10.406.0050 &	n/d	5	2	+	+	+	+	TbRGG2/RRM
Tb11.02.5390	n/d	12	6	+	+	+	+	none
Tb927.4.4160	n/d	12	3	+	+	+	+ /-	none
Tb11.01.8620	n/d	3	4	+	+	+	+	none
Tb927.6.1680	n/d	4	4	+	n/d	+	+	C2H2 ZnF
Tb927.8.8170	n/d	14	4	+	+	n/d	n/d	none
Tb10.389.1910	n/d	4	2	n/d	n/d	+	+	none
Tb927.6.2140	2	7	4	n/d	+	n/d	n/d	hydratase
Tb927.7.800	n/d	9	5	+	n/d	n/d	n/d	none
Tb927.2.1860	n/d	6	7	+	n/d	n/d	n/d	none
Tb11.55.0009	n/d	8	6	n/d	n/d	+	n/d	MRP1

B

Protein ID *T. brucei*	REH2 IP N	REH2 IP 20-30S fracs.	REH2 IP 20-30S fracs. *RNase*	TAP-tagging TbRGG1	TAP-tagging GRBC1 (GAP1)	TAP-tagging GRBC2(GAP2)	Name/motif
Tb10.389.1910	n/d	4	2	n/d	+	+	none
Tb927.3.4030	n/d	8	7	n/d	+	+	none
Tb10.6k15.0310	n/d	2	3	n/d	+	+	none
Tb927.8.2770 &	n/d	7	11	+	n/d	n/d	none

Mass spectrometric analyses of proteins identified in REH2 antibody pulldowns from ~20-30S glycerol gradient fractions, Unique peptides (criteria: two peptides minimum) before or after nuclease treatment. REH2 pull-downs of native editing complexes (N) and REH2 TAP (&) (criteria: one unique peptide minimum). Reported MRB1-immunopurified complex (using mAb43 antibodies) and TAP-purified TbRGG1 and GRBC1/2 complexes. GRBC1/2 (+) indicates proteins found in both GRBC1 and GRBC2 purifications, whereas (+/-) indicates detection only in the GRBC1 purification. Some proteins listed as putative GRBC components by Weng et al. are ribosomal proteins, and were RNase sensitive in the REH2 pulldowns (Fig. S8); (B) RNase resistant REH2 associated proteins that were not considered as subunits of MRB-related complexes, but were found in independent TAP-purifications—Many proteins in a study by Hashimi and collaborators were not present in at least one of three TAP-purifications: TbRGG1, GBP1 and GBP2 (i.e., were no considered as TbRGG1 complex subunits) (21). The identified REH2 associated components correspond to hypothetical proteins with unknown motifs (criteria: two unique peptides). Proteins in native editing complexes and an REH2 TAP-purification (&) (criteria: one unique peptide). Some reported proteins are ribosomal components.

Seven RNase-resistant proteins (out of thirteen) were common to the known MRB-related complexes, namely: REH2, GRBC2, Ribosomal S2 homolog, TbRGG2, and hypothetical proteins Tb11.02.5390, Tb927.4.4160, and Tb11.01.8620 (Table 5A). Likely, GRBC1 is also in this group but it was not detected with the RNase treatment implying that its proposed 1:1 ratio with GRBC2 in GRBC complexes (111) is not strictly conserved in the immuno-purified REH2 complexes. Apparent differences with known MRB-related complexes include five proteins found in at least one of these complexes but not in the REH2 pull-down (Table 6), and fifteen RNase-resistant proteins exclusively found in the REH2 pulldown, including the putative helicase REH1 (mHel61; Table 7) (114) and hypothetical proteins with unrecognized motifs (Table 8).

Table. 6. RNase-sensitive REH2 Interactions and proteins only found in other MRB-related complexes.

	REH2 IP			mAb43 IP or TAP-tagging				
Protein ID	**N**	**20-30S fractions**		**MRB1**	**tbRGG1**	**GRBC1/2**		**Name/motif**
T. brucei			***RNase***				***RNase***	
Tb927.7.2570	5	6	n/d	+	+	+	+	GRBC1
Tb927.3.1820 &	1	2	n/d	n/d	+	n/d	n/d	GRP
Tb11.01.7290	n/d	3	n/d	n/d	+	+	n/d	MERS1/Nudix
Tb927.6.2230 &	n/d	8	n/d	n/d	+	+	n/d	TbRGG1/RGG
Tb927.2.3180	n/d	2	n/d	n/d	n/d	+	+	PPR1/PPR
Tb927.3.1590	n/d	8	n/d	+	n/d	n/d	n/d	ATPase ABC-class
Tb927.2.6070	n/d	5	n/d	+	n/d	n/d	n/d	RBZ Zinc finger
Tb10.6k15.0150	n/d	3	n/d	+	n/d	+	n/d	none
Tb11.01.4860	n/d	n/d	n/d	n/d	n/d	+	-/+	MRP2
Tb11.02.5120	n/d	n/d	n/d	n/d	n/d	+	-/+	PPR repeats
Tb927.8.8180 &	n/d	n/d	n/d	+	n/d	+	+	none
Tb927.6.2180	n/d	n/d	n/d	n/d	n/d	+	+	none
Tb927.3.4920	n/d	n/d	n/d	n/d	+	n/d	n/d	LETM1

REH2-immunopurified proteins from sedimentation fractions that were lost after the RNase/MN treatment and proteins in reported MRB-related complexes but not detected in our REH2 pulldowns (criteria: two unique peptides minimum). Proteins also found in native editing complexes and in our TAP-REH2 purification (&) (criteria: one unique peptide minimum).

Table. 7. REH2-immunopurified proteins that bear conserved motifs associated with RNA processes but were not detected in MRB1, TbRGG1 or GRBC complexes.

Other proteins with motifs associated with RNA processes				
Protein ID	**N**	**20-30S fractions**		
T. brucei			***RNase***	**Name/motif**
Tb11.01.0610	n/d	3	3	mHel61 (REH1)
Tb927.7.3950	n/d	2	n/d	RET1
Tb11.02.5820	n/d	3	n/d	KPAP1/PAP
Tb11.01.3710	n/d	2	n/d	PTB1-regulated
Tb927.4.3020	n/d	6	n/d	DEAH-helicase
Tb11.01.5980*	n/d	2	n/d	PPR

The unique peptide criteria was as above.

Table 8. RNase/MN-resistant REH2-immunopurified proteins that lack known conserved motifs and were not detected in MRB1, TbRGG1 and GRBC1/2 complexes.

Hypothetical prots. without known motifs			
Protein ID	**N**	**20-30S fractions**	
T. brucei			***RNase***
Tb927.5.2930	1	2	4
Tb11.01.0880	n/d	4	2
Tb10.389.1785	n/d	2	2
Tb927.1.1730	n/d	9	3
Tb10.70.6170	n/d	7	5
Tb11.02.4120	n/d	3	4
Tb11.01.2840	n/d	2	2
Tb09.160.0370	n/d	4	5
Tb10.6k15.3400	n/d	2	2
Tb927.1.3010	n/d	7	4

The TbRGG1-associated complex described by Hashimi and collaborators (110) consists of proteins detected in TAP-purifications of all three TbRGG1, GAP1 (GRBC1) and GAP2 (GRBC2). Among proteins from the Hashimi study that were absent in at least one of their purifications, we found four RNase-resistant REH2 associated proteins (Table 5B). These four proteins were not found in MRB1 and GRBC complexes isolated by other labs (111,112) except for Tb10.389.1910 which was present in the latter (see Tables 5A-B). Thus, some proteins may be preferential or even unique

components of REH2 complexes but this has to be confirmed by subsequent cross-tagging/purifications studies. Thus, besides a common core of at least 7-to-8 polypeptides in REH2 and other MRB-related assemblies there may be significant compositional differences among these novel multi-protein particles.

In contrast to the above examined sedimentation fractions, an REH2 pulldown of purified native editing complexes contained only six reported components of MRB-related complexes (including three RNase-resistant interactions). This could reflect differences in REH2 complex composition or relative abundance of REH2 in the samples examined, since the editing complexes had a somewhat smaller number of unique peptides (Table 5). Interestingly, the protein Tb927.5.2930 may represent a REH2-specific cofactor as it was detected in REH2 pull-downs of both native editing complexes and sedimentation fractions, but was not found in other MRB-related complexes (Table 8). In fact, affinity-purification of this protein contained REH2, helicase activity, and RECC subunits (B.R.M. et al., unpublished data). Finally, REH2 co-purified with most known subunits of both RNA editing complexes and mitochondrial ribosomes (53,139), largely via RNase-sensitive associations (Tables 9 and 10).

Table 9. REH2 co-immunopurified with most subunits of editing complexes.

RNA editing complex (core subunits)			
Protein ID	20-30S fractions		name
T. brucei		*RNase*	
Tb10.70.3850	12	7	REX2
Tb10.70.2090	6	2	MP18 (KREPA6)
Tb927.2.2470	7	3	MP81 (KREPA1)
Tb927.7.1550	4	n/d	RET2
Tb09.160.2970	2	n/d	REL1
Tb927.1.3030	3	n/d	REL2
Tb10.70.1580	3	n/d	REN2
Tb10.70.1750	4	n/d	REN3
Tb927.8.620	5	n/d	MP42 (KREPA3)
Tb11.03.0180	4	n/d	MP44 (KREPB5)
Tb11.02.0490	3	n/d	MP46 (KREPB4)
Tb09.160.4130	3	n/d	MP47 (KREPB7)
Tb10.6k15.2310	2	n/d	MP63 (KREPA2)

Criteria: two unique peptides minimum. Additional subunits were found in single-peptide searches. Editing proteins were not detected in mock purifications without antibody or using an "empty" TAP-vector (data not shown). Editing subunits or proteins in the previous tables were not detected in mock purifications without antibody or using an"empty"TAP-vector.

Table. 10. REH2 co-immunopurified with most mitochondrial ribosome subunits.

MITOCHONDRIAL RIBOSOMAL PROTEINS			
Protein ID	20-30S fraction		Name/motif
T. brucei		*RNase*	
Tb11.02.0130	7	2	SSU unique
Tb11.02.3810 &	8	2	TbMRPL4
Tb10.6k15.3290 &	2	2	LSU unique
Tb11.02.2710	14	4	SSU sub rRNA methylase
Tb11.01.6620 &	6	2	TbMRPL38
Tb927.7.4710	2	3	TbMRPL46 human homolog
Tb927.2.4400 &	11	2	SSU unique
Tb11.02.4810	2	5	Oxidoreductase complex
Tb927.3.3660	2	2	Oxidoreductase complex
Tb927.7.3050 &	6	2	SSU unique
Tb927.8.1880	2	2	LSU unique
Tb11.01.2340 &	4	2	LSU unique
Tb927.7.4550	3	n/d	TbMRPL7/12 60S-like
Tb11.01.3500 &	2	n/d	LSU unique
Tb11.01.3300	3	n/d	SSU unique
Tb11.02.1110 &	3	n/d	TbMRPL27
Tb11.01.1840	2	n/d	LSU unique
Tb927.3.770	4	n/d	SSU unique
Tb927.7.7010 &	2	n/d	LSU unique
Tb10.70.0530	7	n/d	TbMRPS5
Tb10.389.0130	3	n/d	TbMRPS8
Tb927.7.3960	2	n/d	TbMRPL16
Tb927.1.1200	5	n/d	TbMRPS15
Tb10.406.0510	3	n/d	TbMRPS11
Tb927.8.5280	3	n/d	TbMRPS34
Tb927.5.3980 &	5	n/d	TbMRPL15
Tb10.61.3110	4	n/d	SSU unique
Tb10.70.4220	2	n/d	SSU unique
Tb927.5.4120	2	n/d	LSU unique
Tb09.160.2250 &	2	n/d	LSU unique
Tb927.6.1440	5	n/d	LSU unique
Tb927.3.820 &	3	n/d	LSU unique
Tb11.02.5670 &	3	n/d	TbMRPS16
Tb09.160.3800	3	n/d	SSU unique
Tb927.4.4600	3	n/d	TbMRPL43
Tb927.3.5610 &	6	n/d	TbMRPL3
Tb927.7.3460	5	n/d	TRP-2
Tb927.5.3410 &	4	n/d	TbMRPL9
Tb927.6.4200 &	4	n/d	DUF 390
Tb927.3.5240 &	6	n/d	DUF436 PPR
Tb10.26.1000	5	n/d	LSU unique
Tb10.6k15.3900	3	n/d	TbMRPS18
Tb927.6.2480	5	n/d	LSU DnaJ
Tb927.5.3110 &	3	n/d	TbMRPL49
Tb09.211.4510	2	n/d	SSU unique
Tb11.02.3180 &	8	n/d	SSU PPR
Tb927.6.4560	4	n/d	SSU unique
Tb11.47.0024	6	n/d	SSU unique
Tb927.3.1710 &	2	n/d	TbMRPL24
Tb927.1.1160	2	n/d	LSU PPR
Tb927.4.4610	3	n/d	LSU unique
Tb927.1.2990 &	8	n/d	SSU PPR
Tb927.3.2260 &	3	n/d	SSU Ubiquitin
Tb927.7.3510	4	n/d	LSU unique
Tb927.8.4550	2	n/d	SSU unique
Tb927.8.3300	5	n/d	LSU unique
Tb11.01.1600	2	n/d	LSU unique
Tb10.70.3840	4	n/d	SSU unique
Tb927.6.1250	3	n/d	TbMRPS29
Tb10.70.7650	4	n/d	TbMRPL29
Tb11.01.1910 &	5	n/d	SSU ECH
Tb10.70.7960	3	n/d	LSU PPR
Tb09.211.2430	2	n/d	SSU unique
Tb927.8.3110	4	n/d	TbMRPS9
Tb927.5.1790	2	n/d	SSU unique
Tb09.211.0810	2	n/d	LSU GTPase
Tb927.5.4040 &	2	n/d	SSU unique
Tb927.7.1640	3	n/d	LSU GTPase
Tb927.5.1510	2	n/d	SSU unique
Tb11.01.3710	2	n/d	SSU ADP-ribosyl-GH
Tb927.8.5200	11	n/d	SSU unique
Tb09.211.4580 &	4	n/d	LSU Lipase_3
Tb927.6.2080	2	n/d	SSU unique
Tb927.5.3870	3	n/d	LSU unique
Tb11.01.7510 &	3	n/d	SSU unique
Tb09.160.5240	2	n/d	TbMRPL47
Tb11.55.0016	5	n/d	LSU unique
Tb11.02.2250 &	2	n/d	TbMRPL52
Tb927.8.4860	2	n/d	SSU PPR_s
Tb927.3.970	2	n/d	SSU NAD binding
Tb927.5.2070	2	n/d	LSU unique
Tb11.01.1600	2	n/d	LSU unique
Tb10.70.3840	4	n/d	SSU unique
Tb927.6.1250	3	n/d	TbMRPS29
Tb10.70.7650	4	n/d	TbMRPL29
Tb11.01.1910 &	5	n/d	SSU ECH

Several subunits were also detected in our REH2 TAP-purification (&). Criteria: two unique peptides minimum. Additional subunits were found in single-peptide searches. Mitochondrial ribosomal proteins were not detected in mock purifications without antibody or using an "empty" TAP-vector.

An association with mitochondrial ribosomes was not implied for other MRB-related complexes, although four ribosomal proteins were inadvertently included in a model of the GRBC complex core (111). Importantly, none of the protein components discussed above were found in mock 182mmune-purifications or using an "empty"-TAP vector. Other RNase-resistant interactions in the pull-downs correspond to common metabolic proteins or proteins that were detected in our mock purifications. Thus, the REH2 RNPs described here and previously reported MRB-related complexes are similar but not identical in composition and apparent RNA-linked interactions with other mitochondrial components.

Discussion

RNA helicases may be the largest group of enzymes in RNA metabolism from bacteria to humans, and are usually assembled in macromolecular RNPs such as spliceosomes and ribosomes. Helicases are often multifunctional within the same cell and understanding the basis of their specificity, particularly the relevant cofactors and substrates, remains a major challenge in biology (187).

REH2 is a novel ~241 kDa factor in kinetoplastids that bears an unusual dual combination of DExH-box helicase and dsRBD motifs only previously observed in the eukaryotic RNA helicase A (264), and to our knowledge it represents the only reported example of a DExH-box helicase in mitochondria besides Suv3p in yeast and its orthologues (265). Recent RNAi studies by Hashimi et al., and in our lab showed that

repression of REH2 decreases RNA editing and total gRNA levels (Figs. 40C-D, and data not shown) (255).

The REH2 complexes we described appear to be novel particles that overlap in composition with recently published "MRB-related complexes" namely, MRB1, TbRGG1 and GBRC complexes. These complexes exhibit compositional differences (Table 5) but they share a few subunits that resisted extensive RNase treatment in REH2 purifications, which we propose form a scaffold core for the assembly of more dynamic proteinprotein and protein-RNA interactions. RECC binds REH2 and GRBC (111) RNPs, but an association with MRB1 and TbRGG1 complexes was not detected, probably reflecting compositional or stoichiometric differences that control this interaction. TbRGG1 itself is bound via RNA (Table 6) (110). Association of REH2 with RECC subunits was observed in REH2 antibody pulldowns of prepurified native editing complexes and in converse immunoprecipitations (Figs. 37 and 39E). Consistent with a transient interaction (a) REH2 was readily detected in native but not in affinity-purified editing complexes, (b) TAP-REH2 purifications showed only a small amount of editing subunits after concentration of the final elutes, and (c) the majority of editing complexes appeared in the unbound fraction of REH2 pulldowns. A small fraction of REH2 consistently associated with editing proteins despite extensive treatments with Mnase and Rnases A, T1 and V1, suggesting that at least some REH2 may directly bind RECC. The majority, however, was clearly RNA bridged (e.g., Fig. 37E and Fig. 39B). Alternatively, a high-affinity RNA linker, not completely removed by protein purification and single-/double-stranded nucleases, could mediate the association with

editing complexes. A TAP-GRBC co-purification with REL1 and RET1 editing subunits was partially sensitivity to RNase A, and thus thought to involve transient contacts (111). GRBC complexes reported by Weng et al. and the REH2 complexes described here contain gRNA, however we also found helicase activity in the REH2 pulldowns. Our data suggest that REH2 is directly responsible for gRNA binding and the helicase activity as both required the dsRBD and a wild-type motif I of REH2 (Fig. 41C). This helicase activity immuno-purified with REH2 from sedimentation fractions containing no visible REH1 (Figs. 41D-F, and data not shown). Interestingly, Missel et al., found a helicase activity of mitochondrial lysates that largely sedimented away from REH1, localized at the top of the gradient (114). Consistent with the motif I mutation effect, the unwinding activity utilized ATP but not a non-hydrolyzable analog, and it dissociated a model pre-mRNA/gRNA hybrid and a continuous helix with 3' but not 5' overhangs. The molecular basis of this substrate selectivity remains to be studied, however our data suggest that REH2 translocates in a 3' to 5' direction from a single-stranded loading region into a helical structure (266). We hypothesize that the dsRBD and DExH-helicase domains of REH2 act in concert to mediate selective binding of properly folded gRNAs. Helical elements of gRNA (267) may be targeted by the 70-amino acid dsRBD, although other REH2 sequences or its co-factors likely contribute to substrate specificity. Subsequent ATP-dependent winding and unwinding remodeling cycles may generate specific gRNA folds with increased affinity for REH2, thereby becoming protected from degradation. Thus, the conformation of gRNAs may be subject to regulation. Consistent with the idea that REH2 and other associated factors

cooperate to provide gRNA-binding specificity, RNAi of REH2 and GRBC (GAP) factors severely decreased the gRNA steady-state levels (Figs. 40C-D) (111,255).

The precise compositional and functional relationship of REH2 complexes with known MRB-related complexes needs to be further studied. The number of RNase-resistant interactions that co-purified with REH2 contrasts with the significantly decreased S value of REH2 after RNase treatments or dsRBD deletion. To reconcile these seemingly discordant observations we propose that REH2 may form multiple particles of variable composition or stoichiometry. In this line of thought, we indicated that related complexes, isolated via e.g., either REH2, GRBC or TbRGG1 proteins share a common scaffold core but differ in multiple dynamic components. Such components could include specificity factors that link these particles with various aspects of mitochondrial gene expression. A putative collection of purified REH2 complexes may include most if not all proteins in Tables 5A-B, and potentially other RNase-resistant components we observed but additional studies are underway to examine this further.

The broad heterodispersion of REH2 in sedimentation gradients likely reflects higher-order assemblies linked by RNA that comprise REH2 RNPs, other MRB-related assemblies, RECC and several factors including KPAP1, MERS1, RET1, MRP and PPR1. The pentatricopeptide protein PPR1 was associated with processing/stability of mitochondrial RNA including editing and ribosomal transcripts, specifically with regulation of long poly (A) tails (137,268,269). Consistent with the above model: (a) these factors were found in REH2 pulldowns before but not after an extensive RNase treatment (Table 6), (b) either RNase or dsRBD deletion decreased the S value of REH2,

and (c) REH2 co-immunoprecipated with RECC subunits in ~20S and >40S fractions (Fig. 39). Importantly, mitochondrial ribosomes were a major component of REH2 purifications via specific antibodies or TAP tagging (but were not found in mock preparations; Table 10), and their association was bridged by RNA. We found a few ribosomal proteins in reported TAP purifications of GRBP, KPAP1, MERS1, MRP, TbRGG1 and editing proteins (110,111), although this had passed unnoticed as the composition of mitochondrial ribosomes was reported after these studies (139). Furthermore, Osato et al. found rRNA in TAP isolations of editing complexes (270).

Collectively data by us and by others labs are consistent with a model whereby transient and dynamic RNA inter-linked networks in mitochondria functionally integrate and coordinate mitochondrial machineries in RNA maturation, stability and translation (schematized in Fig. 43). In this context, REH2 may be multifunctional. Relevant to this model and the integration of ribosomes in particular, is the presence of a S2 ribosomal homolog in the core of REH2 RNPs and other MRB-related complexes (Table 5A; Fig. 43). Interestingly, S2 was not found in isolated mitochondrial ribosomes of *T. brucei* (139), and we speculate that this protein may help to functionally link MRB-related complexes with ribosomes. Finally, the multi component networks we propose may be similar to the L*b complexes recently detected in Leishmania using blue native gel electrophoresis by Osato et al. which the authors proposed represent the active holoenzyme or “editosome” (270).

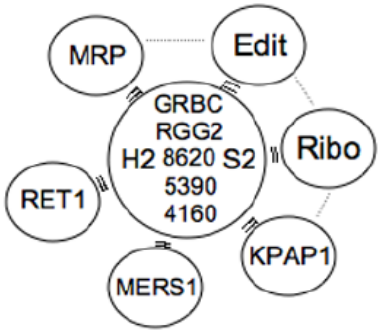

Figure 43. Proposed RNA-interlinked molecular networks in kinetoplastid mitochondria. Machineries and factors for RNA processing, stability and translation that co-purify with REH2 RNPs via RNA, that we propose reflect transient supramolecular assemblies that coordinate mitochondrial gene expression. For simplicity, RNA bridges (double lines) irradiate from the center but alternative or additional RNA and protein contacts could exist across components in the network (dotted lines). The arbitrary center of the model includes seven RNase-resistant proteins in the REH2 pull-down that are common to the known MRB-related complexes. They were abbreviated as follow: REH2 (H2), GRBC2 (GRBC), Ribosomal S2 homolog (S2), TbRGG2 (RGG2), and hypothetical proteins Tb11.01.8620 (8620), Tb11.02.5390 (5390) and Tb927.4.4160 (4160) (Table 5A). At least GRBC1 may also be included in this protein array (see text).

CHAPTER VI

CONCLUSIONS

Summary

The central dogma of molecular biology is one of the greatest intellectual achievements in all of history. In essence, it describes the flow of genetic information from one biological polymer to another (271). When RNA editing was discovered over two decades ago (40), it seemed that trypanosomes were challenging the very basis of molecular biology because it appeared that the information content of RNA was altered seemingly out of nowhere . We now know that Crick's idea remains in full-force as the unifying concept of modern biology. RNA editing, however, has also emerged as a conserved characteristic of RNA metabolism that expands the genetic and functional repertoire of organisms (10).

RNA editing in trypanosomes is a post-transcriptional maturation process in which uridylates are inserted or deleted from mitochondrial mRNAs at specific editing sites by an enzymatic cycle consisting of mRNA cleavage, U-insertion/deletion, and RNA ligation (42). Editing takes pace in the context of the RNA editing core complex, a megadalton multisubunit complex, and is directed by small, trans-acting guide RNA molecules. While great progress has been made in the characterization of the mechanism of RNA editing and the subunit composition of RECC, fundamental questions regarding the mechanism of substrate recognition remain that have profound

implications for RNP assembly, editing site determination and the regulation of the catalytic cycle. In addition, the regulation of RNA editing by its integration with other steps in RNA metabolism is an important area of research that is only beginning to be addressed.

RNA-Protein Interactions in Trypanosome RNA Editing

In Chapter II, we found that at least four proteins interact directly with a model editing site for U-deletion using purified RECC and a synthetic, photoreactive 4-thioU-substituted RNA. These RNA-protein interactions involve RECC subunits since they precisely co-purify with editing activity upon extensive chromatographic purification and are immune-precipitated with monoclonal antibodies specific for RECC subunits.

We hypothesize that the cross-links we observed may play a role in editing site recognition by RECC, since we found that they exhibit a structural selectivity for a double-strand/single-strand junction upstream of an "anchor" duplex between an mRNA and gRNA, a long-recognized hallmark of an editing site (49,50). In fact, base-pairing of the editing site resulted in the inhibition of interactions that allow cross-linking to occur. Furthermore, competition analyses suggest that RECC preferentially associates with a pre-mRNA:gRNA hybrid due to recognition of its features.

In Chapter III, we dissected the functional determinants of a substrate for full-round U-insertion that underscore the importance of RECC-substrate contacts during the

editing process and that may have implications for the mechanism of substrate recognition by RECC (see below).

We found that the sequence composition of helical elements flanking the editing site are not a requirement for efficient editing, implying that they are dispensable for substrate recognition by RECC. In fact, substrate recongnition of substrates outside the internal loop is independent of sequence, with artificial duplexes spanning one helical turn of dsRNA supporting editing.

We hypothesize that duplexes flanking the editing site may be recognized by RECC endonuclease subunits containing conserved RNase III and double-stranded RNA-binding domain, as studies in other systems suggest that RNase III proteins can bind dsRNA spanning one helical turn. Additionally, 2'-hydroxyl substitutions in the mRNA and gRNA strands helped define substrate regions important for recognition and catalysis. Interestingly, a single 2'-deoxy substitution immediately upstream of the editing site completely abolished pre-mRNA cleavage but not association to RECC, suggesting that this residue may be functionally involved in catalysis.

In Chapter IV, we described relatively simple substrate secondary structural determinants for association and guide-RNA directed RNA cleavage, the first enzymatic step of the editing process, using a combination of parallel binding and enzymatic assays. This was the first analysis of direct RNA binding by purified RECC. The ribonucleoproteins we detected by EMSA contained adenylylatable ligases and co-immunoprecipitated with known RECC subunits.

Our results suggest that potential RECC ligands must satisfy an appropriate combination of differentially recognized single-stranded and double-stranded RNA elements for their association and cleavage. Importantly, we found that substrate association and cleavage can be uncoupled. In addition to a requirement for a 5' 12-nucleotide overhang, cleavage requires a duplex of at least 15 base-pairs. RNAs bearing a duplex as short as 11 base-pairs are bound by RECC. The single-stranded 5'-overhang of 12 nucleotides appears to be absolutely required for both association and cleavage.

We also reported the results of a mass-spectrometric analysis of RECC purified in our laboratory which show that, in addition to the majority of reported RECC subunits, we found components of the MRP complex.

In addition, we identified REN1, the RNase III deletion endonuclease, as one of the major cross-links we observed between purified RECC and photoactivatable RNA model substrates (Chapters II and III).

A Model for Substrate Recognition by RECC

Based on the observations above, I propose the following model of substrate recognition by RECC *in vitro* (Fig. 44). First, RECC recognizes helical regions flanking the editing site in a sequence-independent manner, most likely though interactions with 2'-hydroxyls and the phosphodiester backbone. The length of this binding region could minimally be one turn of A-form dsRNA (i.e., 11 bp). RECC proteins (such as REN1, REN2 or REN3) containing RNase III and double-stranded RNA binding domains are

prime candidates to mediate this interaction, as bacterial RNase III enzymes have been found to possess similar binding requirements (145,148,149,152). Additionally, loop residues in mRNA may be recognized by single-stranded RNA-binding proteins primarily via their 2'-hydroxyls and possibly base-specific contacts, and through such interactions, the identity of editing site is established (U-insertion or U-deletion). These check-points, once fulfilled, commit RECC for catalysis.

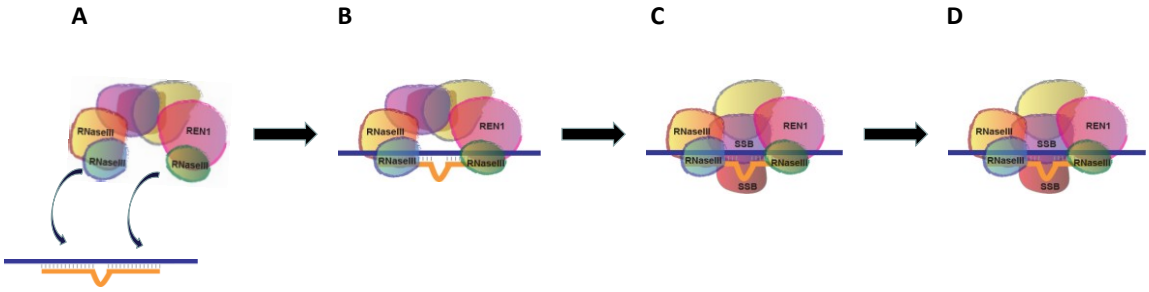

Figure 44. Model for Substrate Recognition by RNA editing core complexes. A, B) RECC targets substrate helical regions flanking the editing site in a sequence-independent manner through interaction of substrate phosphate backbone groups and 2'-hydroxyls with RNase III and dsRBDs domains of RECC proteins. C) Unpaired substrate residues in the internal loop are recognized by single-stranded RNA-binding RECC proteins (such as those containing OB-folds and zinc-finger domains (54,60)). D) Once the determinants that distinguish substrates are satisfied, catalysis begins.

DExH-box Proteins in Kinetoplastid Mitochondria

In Chapter V, we characterized REH2, a novel DExH-box protein that also contains a double-stranded RNA-binding domain. Genetic evidence suggests that this protein plays a role in RNA editing and is required for the stability of guide RNAs (Chapter V and (255).

REH2 forms ribonucleoprotein complexes whose protein composition partially overlaps the recently reported "mitochondrial RNA-binding" or MRB-like complexes (111,254). We propose that the overlapping components, including REH2, form a scaffolding core for the assembly of more dynamic/transient protein-protein and protein-RNA interactions. In addition, REH2 associates with RECC subunits and mitochondrial ribosomes primarily via RNA.

REH2 complexes contained gRNA and displayed RNA unwinding activity *in vitro*. Our evidence suggests that REH2 is directly responsible for binding RNA and for helicase activity as gRNA association and RNA unwinding required the presence of functional double-stranded RNA binding and helicase motifs in REH2. These results led us to suggest a hypothesis whereby REH2 would selectively bind and remodel guide RNAs, generating gRNA conformations with increased affinity for either REH2 or other downstream factors, thereby becoming protected from RNA degradation pathways.

Finally, mass-spectrometic analyses of REH2 complexes revealed links to several protein complexes involved in mitochondrial gene expression and we hypothesize that REH2 or the REH2 complex may link RNA processing and translation machineries.

In conclusion, this dissertation presents *in vitro* investigations aimed at the study of RNA-protein interactions between RECC and RNA ligands as well as the determination of the secondary structural determinants that govern them. Observations from these studies show that show that substrate association, cleavage and full-round editing by purified RECC are directed by hierarchical substrate determinants that increase in complexity as these editing stages progress. These results suggest a model for the recognition of double-stranded RNA editing substrates by RECC (Fig. 10).

In addition, this dissertation also presents the characterization of a novel DExH-box helicase, named REH2, catalogues its macromolecular interactions and suggests a role in RNA editing. REH2 appears to be intimately involved in macromolecular interactions that integrate diverse processes mediating mitochondrial gene expression.

Future Directions

Testing Contributions of Individual RNA-binding Domains to RNA Binding by RECC

In order to test the contributions of a particular protein's domains RNA-binding domains, one could express a TAP-tagged mutant protein in which either the RNA-binding domain is deleted or alternatively, where functionally important residues are mutated and affinity-purify RECC via the tagged mutant. Mutant-containing complexes could then be examined for RNA binding *in vitro*. A major caveat would be that the mutations do not disrupt the integrity of RECC.

Testing the contributions to RNA binding by conserved domains in REN1 would be an important first step, given its functional significance in RNA editing, the fact that it is one of the major cross-links we identified and the predictions of our model.

Identification of Additional RNA-Binding RECC Subunits

In addition to the four RECC proteins that contact functional RNA editing substrates at the editing site, we expect that additional subunits contact the RNA substrate because the majority of RECC subunits contain domains that suggest interaction with nucleic acids (54,60). Thus, identifying RNA-binding RECC subunits may provide information into their role in substrate recognition and their function within the complex.

I propose an approach that is based on RNA-protein cross-linking, affinity purification and mass-spectrometry (272) in order to identify additional RNA-binding

RECC subunits. In this approach, purified RECC would be cross-linked using 260 nm UV light to an RNA ligand that is biotinylated at one end. The cross-linked proteins would then be affinity-purified using streptavidin-coupled magnetic beads under conditions that disrupt RECC interactions in order to remove RECC subunits that are not cross-linked to RNA. The sample affinity-purified sample would then be treated with proteases and analyzed by mass-spectrometry to determine the identity of RECC proteins cross-linked to RNA. Using both functional RNA editing substrates and RNAs that are not expected to be specifically recognized by RECC as baits, this approach could potentially discriminate between RECC subunits that make substrate-specific contacts versus those that are involved in general RNA-binding.

Characterization of the Function of REH2

The data in Chapter V suggest that REH2 may function at various stages in RNA metabolism. The function of REH2 can be investigated by examining the consequences of RNAi of REH2 expression, over-expression of mutant versions, or by gene knock-out.

RNAi of REH2 expression decreases the levels of gRNAs [Chapter V and (255)]. We proposed that REH2 stabilizes gRNAs by folding them into structures that are either refractory to degradation or that are bound by downstream proteins and thereby become protected from degradation. To test this hypothesis, one could induce RNAi of REH2 in combination with treatments that inhibit cellular transcription (such as treatment of cells with actinomycin D) to determine the stability of gRNAs after REH2 is depleted. The stability of gRNAs could be established by determining their half-life in the cell after

transcriptional arrest. If REH2 stabilizes gRNAs, depletion of REH2 would lead to a faster decay of gRNA molecules after transcriptional arrest. Additionally, over-expression of ATPase or RNA-binding mutants can provide information of the contribution of particular protein domains to gRNA stabilization. In our experiments in Chapter V, we detected gRNAs by 5'-labeling with guanylyltransferase in the presence of [α-^{32}P]-GTP. This approach is useful for the detection of bulk gRNA levels; however, it is possible that only a subset of gRNAs is affected by REH2 RNAi. Thus, one could use northern blotting as an alternative approach for guide RNA detection that carries the added benefit of following the fate of a single transcript upon REH2 depletion, or the over-expression of a REH2 mutant.

REH2 interacts with RECC and mitochondrial ribosomes (Chapter V). Thus, it is conceivable that REH2 depletion will affect these processes. In fact, this is the case for editing, a recent report shows that RNA editing levels *in vivo* of a subset of transcripts are disrupted upon REH2 depletion (255). One can envision many potential steps during editing where RNA helicase or RNPase activity may be required, ranging from unfolding stable structures in mRNA to the dissociation of an mRNA-gRNA duplex once editing is accomplished. Therefore it will be important to determine if there are steps during the editing process that are specifically affected by the absence of REH2. This can be done by examining the distribution of unedited, partially-edited, and fully-edited forms for a particular transcript by cloning the products of reverse transcription-PCR amplified using primer pairs targeting transcript regions whose

sequence is not altered by editing. Alternatively, one could examine the effect of addition of REH2 to *in vitro* editing assays.

Examining the effects of REH2 depletion on mitochondrial translation, on the other hand, will be more technically challenging. Conceivably, the role of REH2 in mitochondrial translation could be determined by comparing mitochondrial protein samples of uninduced and induced RNAi cells in a two-dimensional gradient blue-native/SDS-polyacrlyamide gel electrophoretic system as described (43,142,143), or alternatively, by western blotting. Proteins synthesized in mitochondria display anomalous electrophoretic behavior, which could complicate their analysis. A more feasible approach might be to examine the effect of REH2 RNAi on mitochondrial translation indirectly by monitoring the enzymatic activities of mitochondrial complexes that contain proteins subunits synthesized in mitochondria. If REH2 affects mitochondrial translation, it should negatively affect complexes that contain subunits synthesized by the mitochondrial translation machinery and this effect would be reflected in their enzymatic activities.

Identification of RNAs that Associate with REH2

REH2 complexes bind gRNA (Chapter V), however, it is unclear if REH2 associates with this class of RNA exclusively. Although the RNA-dependent interaction of REH2 with RECC could be mediated by gRNA, the interaction of REH2 with mitochondrial ribosomes implies that REH2 may associate with other classes of RNAs. The association of REH2 with editing and translation machineries suggests that REH2

may serve as a link between the two processes. Alternatively, REH2 may play distinct functions at the editing and translation stages. Thus, determining the identity of RNAs that associate with REH2 may provide insight into the biological function of REH2.

To identify cellular RNAs associated with REH2, one could use CLIP (cross-linking followed by immunoprecipitation (273,274). Briefly, RNA-protein complexes are cross-linked in situ by the irradiation of intact cells with UV light. The protein of interest (along with RNAs cross-linked to it) is immunoprecipitated from extracts made from the irradiated cells using specific antibodies and purified by SDS-PAGE. Following the ligation of adaptors, RNAs can be detected by either conventional RT-PCR or deep-sequencing technologies. If this method is successful, it would be interesting to analyse REH2 mutants to characterize the RNAs and processes that are impaired by mutation of REH2 domains.

Characterization of REH3

REH3 is a DExH-box protein that shows significant sequence homology and similar domain architecture to REH2 but it shows interesting functional differences. We hypothesize that REH2 and REH3 may be functionally complementary, at least partially, because of their similarities in sequence and domain structure. Thus the characterization of REH3 may also provide information of the role of REH2 in RNA metabolism in trypanosome mitochondria and is currently ongoing in the lab.

REFERENCES

1. Simpson, A. G., Stevens, J. R., and Lukes, J. (2006) *Trends Parasitol* **22**, 168-174
2. Simpson, A. G., Lukes, J., and Roger, A. J. (2002) *Mol Biol Evol* **19**, 2071-2083
3. Nussbaum, K., Honek, J., Cadmus, C. M., and Efferth, T. (2010) *Curr Med Chem* **17**, 1594-1617
4. Steverding, D. (2008) *Parasit Vectors* **1**, 3
5. Stuart, K., Brun, R., Croft, S., Fairlamb, A., Gurtler, R. E., McKerrow, J., Reed, S., and Tarleton, R. (2008) *J Clin Invest* **118**, 1301-1310
6. Coura, J. R., and Borges-Pereira, J. (2010) *Acta Trop* **115**, 5-13
7. Haile, S., and Papadopoulou, B. (2007) *Curr Opin Microbiol* **10**, 569-577
8. Matthews, K. R. (2005) *J Cell Sci* **118**, 283-290
9. Ochsenreiter, T., and Hajduk, S. (2008) The Function of RNA Editing in Trypanosomes. in *RNA Editing* (Göringer, H. U. ed.), Springer-Verlag Berlin Heidelberg. pp 181-197
10. Gott, J. M., and Emeson, R. B. (2000) *Annu Rev Genet* **34**, 499-531
11. Lukes, J., Hashimi, H., and Zikova, A. (2005) *Curr Genet* **48**, 277-299
12. Ochsenreiter, T., Cipriano, M., and Hajduk, S. L. (2008) *PLoS One* **3**, e1566
13. Farajollahi, S., and Maas, S. (2010) *Trends Genet* **26**, 221-230
14. Jacques, J. P., and Kolakofsky, D. (1991) *Genes Dev* **5**, 707-713
15. Hausmann, S., Garcin, D., Delenda, C., and Kolakofsky, D. (1999) *J Virol* **73**, 5568-5576
16. Blanc, V., and Davidson, N. O. (2003) *J Biol Chem* **278**, 1395-1398
17. Chateigner-Boutin, A. L., and Small, I. (2010) *RNA Biol* **7**, 213-219
18. Stern, D. B., Goldschmidt-Clermont, M., and Hanson, M. R. (2010) *Annu Rev Plant Biol* **61**, 125-155
19. Shikanai, T. (2006) *Cell Mol Life Sci* **63**, 698-708
20. Nishikura, K. (2010) *Annu Rev Biochem* **79**, 321-349
21. Blow, M. J., Grocock, R. J., van Dongen, S., Enright, A. J., Dicks, E., Futreal, P. A., Wooster, R., and Stratton, M. R. (2006) *Genome Biol* **7**, R27
22. Maas, S., Rich, A., and Nishikura, K. (2003) *J Biol Chem* **278**, 1391-1394
23. Levanon, E. Y., Eisenberg, E., Yelin, R., Nemzer, S., Hallegger, M., Shemesh, R., Fligelman, Z. Y., Shoshan, A., Pollock, S. R., Sztybel, D., Olshansky, M., Rechavi, G., and Jantsch, M. F. (2004) *Nat Biotechnol* **22**, 1001-1005
24. Hundley, H. A., and Bass, B. L. (2010) *Trends Biochem Sci* **35**, 377-383
25. Lohan, A. J., and Gray, M. W. (2007) *Methods Enzymol* **424**, 223-242
26. Schuster, J., and Mörl, M. (2004) Mitochondrial tRNA editing. in *Mitochondrial Function and Biogenetics* (Köhler, C., and Bauer, M. F. eds.), Springer-Verlag, Berlin Heidelberg. pp
27. Wolf, J., Gerber, A. P., and Keller, W. (2002) *EMBO J* **21**, 3841-3851
28. Price, D. H., and Gray, M. W. (1999) *RNA* **5**, 302-317

29. Antes, T., Costandy, H., Mahendran, R., Spottswood, M., and Miller, D. (1998) *Mol Cell Biol* **18**, 7521-7527
30. Borner, G. V., Morl, M., Janke, A., and Paabo, S. (1996) *EMBO J* **15**, 5949-5957
31. Marechal-Drouard, L., Kumar, R., Remacle, C., and Small, I. (1996) *Nucleic Acids Res* **24**, 3229-3234
32. Apirion, D., and Miczak, A. (1993) *Bioessays* **15**, 113-120
33. Keegan, L. P., Leroy, A., Sproul, D., and O'Connell, M. A. (2004) *Genome Biol* **5**, 209
34. Gerber, A. P., and Keller, W. (2001) *Trends Biochem Sci* **26**, 376-384
35. Mahendran, R., Spottswood, M. S., Ghate, A., Ling, M. L., Jeng, K., and Miller, D. L. (1994) *EMBO J* **13**, 232-240
36. Mahendran, R., Spottswood, M. R., and Miller, D. L. (1991) *Nature* **349**, 434-438
37. Yang, W., Chendrimada, T. P., Wang, Q., Higuchi, M., Seeburg, P. H., Shiekhattar, R., and Nishikura, K. (2006) *Nat Struct Mol Biol* **13**, 13-21
38. Kawahara, Y., Zinshteyn, B., Chendrimada, T. P., Shiekhattar, R., and Nishikura, K. (2007) *EMBO Rep* **8**, 763-769
39. Scadden, A. D., and Smith, C. W. (2001) *EMBO Rep* **2**, 1107-1111
40. Benne R, B. J., Brakenhoff JPJ, Sloof, and P, B. J., et al. (1986) *Cell* **46**, 819-826
41. Feagin, J. E., Abraham, J. M., and Stuart, K. (1988) *Cell* **53**, 413-422
42. Cruz-Reyes J., and Hernandez, A. (2008) Protein-protein and RNA-protein interactions in U-insertion/deletion RNA editing complexes. in *RNA and DNA Editing* (Smith, H. C. ed.), John Wiley & Sons, Inc., New Jersey. pp 71–98
43. Horvath, A., Berry, E. A., and Maslov, D. A. (2000) *Science* **287**, 1639-1640
44. Ochsenreiter, T., Anderson, S., Wood, Z. A., and Hajduk, S. L. (2008) *Mol Cell Biol* **28**, 5595-5604
45. Speijer, D. (2006) *IUBMB Life* **58**, 91-96
46. Panigrahi, A. K., Schnaufer, A., Ernst, N. L., Wang, B., Carmean, N., Salavati, R., and Stuart, K. (2003) *RNA* **9**, 484-492
47. Aphasizhev, R., Aphasizheva, I., Nelson, R. E., Gao, G., Simpson, A. M., Kang, X., Falick, A. M., Sbicego, S., and Simpson, L. (2003) *EMBO J* **22**, 913-924
48. Cruz-Reyes, J., and Sollner-Webb, B. (1996) *Proc Natl Acad Sci U S A* **93**, 8901-8906
49. Kable, M. L., Seiwert, S. D., Heidmann, S., and Stuart, K. (1996) *Science* **273**, 1189-1195
50. Seiwert, S. D., Heidmann, S., and Stuart, K. (1996) *Cell* **84**, 831-841
51. Cruz-Reyes, J., Zhelonkina, A., Rusche, L., and Sollner-Webb, B. (2001) *Mol Cell Biol* **21**, 884-892
52. Carnes, J., and Stuart, K. (2008) Working Together: the RNA Editing Machinery in *Trypanosoma brucei*. in *RNA Editing* (Göringer, H. U. ed.), Springer-Verlag, Berlin Heidelberg. pp 143-164
53. Panigrahi, A. K., Schnaufer, A., and Stuart, K. D. (2007) *Methods Enzymol* **424**, 3-24

54. Stuart, K. D., Schnaufer, A., Ernst, N. L., and Panigrahi, A. K. (2005) *Trends Biochem Sci* **30**, 97-105
55. Simpson, L., Aphasizhev, R., Gao, G., and Kang, X. (2004) *RNA* **10**, 159-170
56. Simpson, L., Aphasizhev, R., Lukes, J., and Cruz-Reyes, J. (2010) *Protist* **161**, 2-6
57. Carnes, J., Trotter, J. R., Peltan, A., Fleck, M., and Stuart, K. (2008) *Mol Cell Biol* **28**, 122-130
58. Trotter, J. R., Ernst, N. L., Carnes, J., Panicucci, B., and Stuart, K. (2005) *Mol Cell* **20**, 403-412
59. Carnes, J., Trotter, J. R., Ernst, N. L., Steinberg, A., and Stuart, K. (2005) *Proc Natl Acad Sci U S A* **102**, 16614-16619
60. Worthey, E. A., Schnaufer, A., Mian, I. S., Stuart, K., and Salavati, R. (2003) *Nucleic Acids Res* **31**, 6392-6408
61. Panigrahi, A. K., Ernst, N. L., Domingo, G. J., Fleck, M., Salavati, R., and Stuart, K. D. (2006) *RNA* **12**, 1038-1049
62. Rogers, K., Gao, G., and Simpson, L. (2007) *J Biol Chem* **282**, 29073-29080
63. Kang, X., Rogers, K., Gao, G., Falick, A. M., Zhou, S., and Simpson, L. (2005) *Proc Natl Acad Sci U S A* **102**, 1017-1022
64. Ernst, N. L., Panicucci, B., Carnes, J., and Stuart, K. (2009) *RNA* **15**, 947-957
65. Ernst, N. L., Panicucci, B., Igo, R. P., Jr., Panigrahi, A. K., Salavati, R., and Stuart, K. (2003) *Mol Cell* **11**, 1525-1536
66. Aphasizheva, I., and Aphasizhev, R. (2010) *Mol Cell Biol* **30**, 1555-1567
67. Schnaufer, A., Ernst, N. L., Palazzo, S. S., O'Rear, J., Salavati, R., and Stuart, K. (2003) *Mol Cell* **12**, 307-319
68. Schnaufer, A., Panigrahi, A. K., Panicucci, B., Igo, R. P., Jr., Wirtz, E., Salavati, R., and Stuart, K. (2001) *Science* **291**, 2159-2162
69. Kang, X., Gao, G., Rogers, K., Falick, A. M., Zhou, S., and Simpson, L. (2006) *Proc Natl Acad Sci U S A* **103**, 13944-13949
70. Gao, G., Simpson, A. M., Kang, X., Rogers, K., Nebohacova, M., Li, F., and Simpson, L. (2005) *Proc Natl Acad Sci U S A* **102**, 4712-4717
71. Law, J. A., Huang, C. E., O'Hearn, S. F., and Sollner-Webb, B. (2005) *Mol Cell Biol* **25**, 2785-2794
72. Gao, G., and Simpson, L. (2003) *J Biol Chem* **278**, 27570-27574
73. Cruz-Reyes, J., Zhelonkina, A. G., Huang, C. E., and Sollner-Webb, B. (2002) *Mol Cell Biol* **22**, 4652-4660
74. Huang, C. E., Cruz-Reyes, J., Zhelonkina, A. G., O'Hearn, S., Wirtz, E., and Sollner-Webb, B. (2001) *EMBO J* **20**, 4694-4703
75. Palazzo, S. S., Panigrahi, A. K., Igo, R. P., Salavati, R., and Stuart, K. (2003) *Mol Biochem Parasitol* **127**, 161-167
76. Suck, D. (1997) *Nat Struct Biol* **4**, 161-165
77. Kang, X., Falick, A. M., Nelson, R. E., Gao, G., Rogers, K., Aphasizhev, R., and Simpson, L. (2004) *J Biol Chem* **279**, 3893-3899
78. Guo, X., Ernst, N. L., Carnes, J., and Stuart, K. D. (2010) *PLoS One* **5**, e8913

79. Babbarwal, V. K., Fleck, M., Ernst, N. L., Schnaufer, A., and Stuart, K. (2007) *RNA* **13**, 737-744
80. Wang, B., Ernst, N. L., Palazzo, S. S., Panigrahi, A. K., Salavati, R., and Stuart, K. (2003) *Eukaryot Cell* **2**, 578-587
81. Cruz-Reyes, J., Rusche, L. N., Piller, K. J., and Sollner-Webb, B. (1998) *Mol Cell* **1**, 401-409
82. Cruz-Reyes, J., Rusche, L. N., and Sollner-Webb, B. (1998) *Nucleic Acids Res* **26**, 3634-3639
83. Igo, R. P., Jr., Palazzo, S. S., Burgess, M. L., Panigrahi, A. K., and Stuart, K. (2000) *Mol Cell Biol* **20**, 8447-8457
84. Igo, R. P., Jr., Weston, D. S., Ernst, N. L., Panigrahi, A. K., Salavati, R., and Stuart, K. (2002) *Eukaryot Cell* **1**, 112-118
85. Carnes, J., and Stuart, K. D. (2007) *Methods Enzymol* **424**, 25-54
86. Ringpis, G. E., Stagno, J., and Aphasizhev, R. (2010) *J Mol Biol* **399**, 696-706
87. Aphasizhev, R., Aphasizheva, I., and Simpson, L. (2003) *P Natl Acad Sci USA* **100**, 10617-10622
88. Cruz-Reyes, J. (2007) *Methods Enzymol* **424**, 107-125
89. Sollner-Webb, B., Rusche, L. N., and Cruz-Reyes, J. (2001) *Methods Enzymol* **341**, 154-174
90. Alatortsev, V. S., Cruz-Reyes, J., Zhelonkina, A. G., and Sollner-Webb, B. (2008) *Mol Cell Biol* **28**, 2437-2445
91. Cifuentes-Rojas, C., Halbig, K., Sacharidou, A., De Nova-Ocampo, M., and Cruz-Reyes, J. (2005) *Nucleic Acids Res* **33**, 6610-6620
92. Hernandez, A., Panigrahi, A., Cifuentes-Rojas, C., Sacharidou, A., Stuart, K., and Cruz-Reyes, J. (2008) *J Mol Biol* **381**, 35-48
93. Igo, R. P., Jr., Lawson, S. D., and Stuart, K. (2002) *Mol Cell Biol* **22**, 1567-1576
94. Sacharidou, A., Cifuentes-Rojas, C., Halbig, K., Hernandez, A., Dangott, L. J., De Nova-Ocampo, M., and Cruz-Reyes, J. (2006) *RNA* **12**, 1219-1228
95. Cifuentes-Rojas, C., Pavia, P., Hernandez, A., Osterwisch, D., Puerta, C., and Cruz-Reyes, J. (2007) *J Biol Chem* **282**, 4265-4276
96. Deng, J., Ernst, N. L., Turley, S., Stuart, K. D., and Hol, W. G. (2005) *EMBO J* **24**, 4007-4017
97. Zhelonkina, A. G., O'Hearn, S. F., Law, J. A., Cruz-Reyes, J., Huang, C. E., Alatortsev, V. S., and Sollner-Webb, B. (2006) *RNA* **12**, 476-487
98. Schnaufer, A., Wu, M., Park, Y. J., Nakai, T., Deng, J., Proff, R., Hol, W. G., and Stuart, K. D. (2010) *J Biol Chem* **285**, 5282-5295
99. Schumacher, M. A., Karamooz, E., Zikova, A., Trantirek, L., and Lukes, J. (2006) *Cell* **126**, 701-711
100. Vondruskova, E., van den Burg, J., Zikova, A., Ernst, N. L., Stuart, K., Benne, R., and Lukes, J. (2005) *J Biol Chem* **280**, 2429-2438
101. Margossian, S. P., and Butow, R. A. (1996) *Trends Biochem Sci* **21**, 392-396
102. Fisk, J. C., Presnyak, V., Ammerman, M. L., and Read, L. K. (2009) *Mol Cell Biol* **29**, 5214-5225
103. Pelletier, M., and Read, L. K. (2003) *RNA* **9**, 457-468

104. Miller, M. M., Halbig, K., Cruz-Reyes, J., and Read, L. K. (2006) *RNA* **12**, 1292-1303
105. Ammerman, M. L., Fisk, J. C., and Read, L. K. (2008) *RNA* **14**, 1069-1080
106. Goulah, C. C., and Read, L. K. (2007) *J Biol Chem* **282**, 7181-7190
107. Madison-Antenucci, S., Sabatini, R. S., Pollard, V. W., and Hajduk, S. L. (1998) *EMBO J* **17**, 6368-6376
108. Hans, J., Hajduk, S. L., and Madison-Antenucci, S. (2007) *RNA* **13**, 881-889
109. Hernandez, A., Madina, B. R., Ro, K., Wohlschlegel, J. A., Willard, B., Kinter, M. T., and Cruz-Reyes, J. (2010) *J Biol Chem* **285**, 1220-1228
110. Hashimi, H., Zikova, A., Panigrahi, A. K., Stuart, K. D., and Lukes, J. (2008) *RNA* **14**, 970-980
111. Weng, J., Aphasizheva, I., Etheridge, R. D., Huang, L., Wang, X., Falick, A. M., and Aphasizhev, R. (2008) *Mol Cell* **32**, 198-209
112. Panigrahi, A. K., Zikova, A., Dalley, R. A., Acestor, N., Ogata, Y., Anupama, A., Myler, P. J., and Stuart, K. D. (2008) *Mol Cell Proteomics* **7**, 534-545
113. Jankowsky, E., and Bowers, H. (2006) *Nucleic Acids Res* **34**, 4181-4188
114. Missel, A., Souza, A. E., Norskau, G., and Goringer, H. U. (1997) *Mol Cell Biol* **17**, 4895-4903
115. Halbig, K., De Nova-Ocampo, M., and Cruz-Reyes, J. (2004) *RNA* **10**, 914-920
116. Koslowsky, D. J., Riley, G. R., Feagin, J. E., and Stuart, K. (1992) *Mol Cell Biol* **12**, 2043-2049
117. Yamanaka, S., Poksay, K. S., Arnold, K. S., and Innerarity, T. L. (1997) *Genes Dev* **11**, 321-333
118. Bass, B. L. (2002) *Annu Rev Biochem* **71**, 817-846
119. Phan, V., Thomas, A., Malasri, K., and Sutter, C. H. (2010) *Int J Bioinform Res Appl* **6**, 21-36
120. Hausmann, S., Garcin, D., Morel, A. S., and Kolakofsky, D. (1999) *J Virol* **73**, 343-351
121. Thomas, S. M., Lamb, R. A., and Paterson, R. G. (1988) *Cell* **54**, 891-902
122. Rundquist, B. A., and Gott, J. M. (1995) *Mol Gen Genet* **247**, 306-311
123. Gott, J. M., Visomirski, L. M., and Hunter, J. L. (1993) *J Biol Chem* **268**, 25483-25486
124. Shapiro, T. A. (1993) *Proc Natl Acad Sci U S A* **90**, 7809-7813
125. Lukes, J., Guilbride, D. L., Votypka, J., Zikova, A., Benne, R., and Englund, P. T. (2002) *Eukaryot Cell* **1**, 495-502
126. Simpson, A. G., Gill, E. E., Callahan, H. A., Litaker, R. W., and Roger, A. J. (2004) *Protist* **155**, 407-422
127. Grams, J., Morris, J. C., Drew, M. E., Wang, Z., Englund, P. T., and Hajduk, S. L. (2002) *J Biol Chem* **277**, 16952-16959
128. Etheridge, R. D., Aphasizheva, I., Gershon, P. D., and Aphasizhev, R. (2008) *EMBO J* **27**, 1596-1608
129. Koslowsky, D. J., and Yahampath, G. (1997) *Mol Biochem Parasitol* **90**, 81-94
130. Kao, C. Y., and Read, L. K. (2005) *Mol Cell Biol* **25**, 1634-1644

131. Ryan, C. M., Militello, K. T., and Read, L. K. (2003) *J Biol Chem* **278**, 32753-32762
132. Temperley, R. J., Seneca, S. H., Tonska, K., Bartnik, E., Bindoff, L. A., Lightowlers, R. N., and Chrzanowska-Lightowlers, Z. M. (2003) *Hum Mol Genet* **12**, 2341-2348
133. Carpousis, A. J. (2007) *Annu Rev Microbiol* **61**, 71-87
134. Grams, J., McManus, M. T., and Hajduk, S. L. (2000) *EMBO J* **19**, 5525-5532
135. Pollard, V. W., Rohrer, S. P., Michelotti, E. F., Hancock, K., and Hajduk, S. L. (1990) *Cell* **63**, 783-790
136. Clement, S. L., Mingler, M. K., and Koslowsky, D. J. (2004) *Eukaryot Cell* **3**, 862-869
137. Koslowsky, D. J. (2009) *Trends Parasitol* **25**, 252-255
138. Sharma, M. R., Booth, T. M., Simpson, L., Maslov, D. A., and Agrawal, R. K. (2009) *Proc Natl Acad Sci U S A* **106**, 9637-9642
139. Zikova, A., Panigrahi, A. K., Dalley, R. A., Acestor, N., Anupama, A., Ogata, Y., Myler, P. J., and Stuart, K. (2008) *Mol Cell Proteomics* **7**, 1286-1296
140. Maslov, D. A., Spremulli, L. L., Sharma, M. R., Bhargava, K., Grasso, D., Falick, A. M., Agrawal, R. K., Parker, C. E., and Simpson, L. (2007) *Mol Biochem Parasitol* **152**, 203-212
141. Nabholz, C. E., Speijer, D., and Schneider, A. (1999) *Parasitol Res* **85**, 779-782
142. Horvath, A., Nebohacova, M., Lukes, J., and Maslov, D. A. (2002) *J Biol Chem* **277**, 7222-7230
143. Horvath, A., Kingan, T. G., and Maslov, D. A. (2000) *J Biol Chem* **275**, 17160-17165
144. Nicholson, A. W. (1999) *FEMS Microbiol Rev* **23**, 371-390
145. MacRae, I. J., and Doudna, J. A. (2007) *Curr Opin Struct Biol* **17**, 138-145
146. Drider, D., and Condon, C. (2004) *J Mol Microbiol Biotechnol* **8**, 195-200
147. Lamontagne, B., Larose, S., Boulanger, J., and Elela, S. A. (2001) *Curr Issues Mol Biol* **3**, 71-78
148. Ji, X. (2008) *Curr Top Microbiol Immunol* **320**, 99-116
149. Gan, J., Tropea, J. E., Austin, B. P., Court, D. L., Waugh, D. S., and Ji, X. (2006) *Cell* **124**, 355-366
150. Sun, W., Pertzev, A., and Nicholson, A. W. (2005) *Nucleic Acids Res* **33**, 807-815
151. Campbell, F. E., Jr., Cassano, A. G., Anderson, V. E., and Harris, M. E. (2002) *J Mol Biol* **317**, 21-40
152. Blaszczyk, J., Tropea, J. E., Bubunenko, M., Routzahn, K. M., Waugh, D. S., Court, D. L., and Ji, X. (2001) *Structure* **9**, 1225-1236
153. Westphal, H., and Crouch, R. J. (1975) *Proc Natl Acad Sci U S A* **72**, 3077-3081
154. Regnier, P., and Grunberg-Manago, M. (1989) *J Mol Biol* **210**, 293-302
155. Beran, R. K., and Simons, R. W. (2001) *Mol Microbiol* **39**, 112-125
156. Aristarkhov, A., Mikulskis, A., Belasco, J. G., and Lin, E. C. (1996) *J Bacteriol* **178**, 4327-4332

157. Huntzinger, E., Boisset, S., Saveanu, C., Benito, Y., Geissmann, T., Namane, A., Lina, G., Etienne, J., Ehresmann, B., Ehresmann, C., Jacquier, A., Vandenesch, F., and Romby, P. (2005) *EMBO J* **24**, 824-835
158. Dunn, J. J. (1976) *J Biol Chem* **251**, 3807-3814
159. Nagel, R., and Ares, M., Jr. (2000) *RNA* **6**, 1142-1156
160. Pertzev, A. V., and Nicholson, A. W. (2006) *Nucleic Acids Res* **34**, 3708-3721
161. Ge, D., Lamontagne, B., and Elela, S. A. (2005) *Curr Biol* **15**, 140-145
162. Chanfreau, G., Buckle, M., and Jacquier, A. (2000) *Proc Natl Acad Sci U S A* **97**, 3142-3147
163. Chanfreau, G., Legrain, P., and Jacquier, A. (1998) *J Mol Biol* **284**, 975-988
164. Chanfreau, G., Rotondo, G., Legrain, P., and Jacquier, A. (1998) *EMBO J* **17**, 3726-3737
165. Qu, L. H., Henras, A., Lu, Y. J., Zhou, H., Zhou, W. X., Zhu, Y. Q., Zhao, J., Henry, Y., Caizergues-Ferrer, M., and Bachellerie, J. P. (1999) *Mol Cell Biol* **19**, 1144-1158
166. Sun, W., and Nicholson, A. W. (2001) *Biochemistry* **40**, 5102-5110
167. Meng, W., and Nicholson, A. W. (2008) *Biochem J* **410**, 39-48
168. Nicholson, A. W. (1992) *Biochim Biophys Acta* **1129**, 318-322
169. Wu, H., Xu, H., Miraglia, L. J., and Crooke, S. T. (2000) *J Biol Chem* **275**, 36957-36965
170. Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Radmark, O., Kim, S., and Kim, V. N. (2003) *Nature* **425**, 415-419
171. Grishok, A., Pasquinelli, A. E., Conte, D., Li, N., Parrish, S., Ha, I., Baillie, D. L., Fire, A., Ruvkun, G., and Mello, C. C. (2001) *Cell* **106**, 23-34
172. Hutvagner, G., McLachlan, J., Pasquinelli, A. E., Balint, E., Tuschl, T., and Zamore, P. D. (2001) *Science* **293**, 834-838
173. Ketting, R. F., Fischer, S. E., Bernstein, E., Sijen, T., Hannon, G. J., and Plasterk, R. H. (2001) *Genes Dev* **15**, 2654-2659
174. Han, J., Lee, Y., Yeom, K. H., Nam, J. W., Heo, I., Rhee, J. K., Sohn, S. Y., Cho, Y., Zhang, B. T., and Kim, V. N. (2006) *Cell* **125**, 887-901
175. Gregory, R. I., Yan, K. P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N., and Shiekhattar, R. (2004) *Nature* **432**, 235-240
176. Denli, A. M., Tops, B. B., Plasterk, R. H., Ketting, R. F., and Hannon, G. J. (2004) *Nature* **432**, 231-235
177. Bernstein, E., Caudy, A. A., Hammond, S. M., and Hannon, G. J. (2001) *Nature* **409**, 363-366
178. Ma, J. B., Ye, K., and Patel, D. J. (2004) *Nature* **429**, 318-322
179. Song, J. J., Liu, J., Tolia, N. H., Schneiderman, J., Smith, S. K., Martienssen, R. A., Hannon, G. J., and Joshua-Tor, L. (2003) *Nat Struct Biol* **10**, 1026-1032
180. Yan, K. S., Yan, S., Farooq, A., Han, A., Zeng, L., and Zhou, M. M. (2003) *Nature* **426**, 468-474
181. Lingel, A., Simon, B., Izaurralde, E., and Sattler, M. (2003) *Nature* **426**, 465-469
182. Macrae IJ, Z. K., Li F, Repic A, Brooks AN, Cande WZ, Adams PD, Doudna JA. (2006) *Science* **311**, 195-198

183. Bleichert, F., and Baserga, S. J. (2007) *Mol Cell* **27**, 339-352
184. Singleton, M. R., Dillingham, M. S., and Wigley, D. B. (2007) *Annu Rev Biochem* **76**, 23-50
185. Jankowsky, E., and Fairman-Williams, M. E. (2010) An Introduction to RNA Helicases: Superfamilies, Families, and Major Themes. in *RNA Helicases* (Jankowsky, E. ed.), Royal Society of Chemistry. pp 1-31
186. Jankowsky, E. (2010) *Trends Biochem Sci*
187. Jankowsky, E. (2007) *Nature* **449**, 999-1000
188. Jankowsky, E., Gross, C. H., Shuman, S., and Pyle, A. M. (2001) *Science* **291**, 121-125
189. Jankowsky, E., and Jankowsky, A. (2000) *Nucleic Acids Research* **28**, 333-334
190. Pyle, A. M. (2008) *Annu Rev Biochem* **37**, 317-336
191. Fairman-Williams, M. E., Guenther, U.-P., Jankowsky, E. (2010) *Curr Opin Struct Biol* **20**, 313–324
192. Linder, P. (2006) *Nucleic Acids Res* **34**, 4168-4180
193. Le Hir, H., and Andersen, G. R. (2008) *Curr Opin Struct Biol* **18**, 112-119
194. Schwer, B. (2008) *Mol Cell* **30**, 743-754
195. Tange, T. O., Nott, A., and Moore, M. J. (2004) *Curr Opin Cell Biol* **16**, 279-284
196. Nielsen, K. H., Chamieh, H., Andersen, C. B., Fredslund, F., Hamborg, K., Le Hir, H., and Andersen, G. R. (2009) *RNA* **15**, 67-75
197. Andersen, C. B., Ballut, L., Johansen, J. S., Chamieh, H., Nielsen, K. H., Oliveira, C. L., Pedersen, J. S., Seraphin, B., Le Hir, H., and Andersen, G. R. (2006) *Science* **313**, 1968-1972
198. Alcazar-Roman, A. R., Tran, E. J., Guo, S., and Wente, S. R. (2006) *Nat Cell Biol* **8**, 711-716
199. Weirich, C. S., Erzberger, J. P., Flick, J. S., Berger, J. M., Thorner, J., and Weis, K. (2006) *Nat Cell Biol* **8**, 668-676
200. von Moeller, H. e. a. (2009) *Nat Struct Biol* **16**, 247-254
201. Chang, J. H., Cho, Y. H., Sohn, S. Y., Choi, J. M., Kim, A., Kim, Y. C., Jang, S. K., and Cho, Y. (2009) *Proc Natl Acad Sci U S A* **106**, 3148-3153
202. LaRonde-LeBlanc, N., Santhanam, A. N., Baker, A. R., Wlodawer, A., and Colburn, N. H. (2007) *Mol Cell Biol* **27**, 147-156
203. Rusche, L. N., Cruz-Reyes, J., Piller, K. J., and Sollner-Webb, B. (1997) *EMBO J* **16**, 4069-4081
204. Panigrahi, A. K., Schnaufer, A., Carmean, N., Igo, R. P., Jr., Gygi, S. P., Ernst, N. L., Palazzo, S. S., Weston, D. S., Aebersold, R., Salavati, R., and Stuart, K. D. (2001) *Mol Cell Biol* **21**, 6833-6840
205. McManus, M. T., Shimamura, M., Grams, J., and Hajduk, S. L. (2001) *RNA* **7**, 167-175
206. Rusche, L. N., Huang, C. E., Piller, K. J., Hemann, M., Wirtz, E., and Sollner-Webb, B. (2001) *Mol Cell Biol* **21**, 979-989
207. Brecht, M., Niemann, M., Schluter, E., Muller, U. F., Stuart, K., and Goringer, H. U. (2005) *Mol Cell* **17**, 621-630
208. Madison-Antenucci, S., Grams, J., and Hajduk, S. L. (2002) *Cell* **108**, 435-438

209. Blom, D., Burg, J., Breek, C. K., Speijer, D., Muijsers, A. O., and Benne, R. (2001) *Nucleic Acids Res* **29**, 2950-2962
210. Muller, U. F., Lambert, L., and Goringer, H. U. (2001) *EMBO J* **20**, 1394-1404
211. Aphasizhev, R., Aphasizheva, I., Nelson, R. E., and Simpson, L. (2003) *RNA* **9**, 62-76
212. Vanhamme, L., Perez-Morga, D., Marchal, C., Speijer, D., Lambert, L., Geuskens, M., Alexandre, S., Ismaili, N., Goringer, U., Benne, R., and Pays, E. (1998) *J Biol Chem* **273**, 21825-21833
213. Allen, T. E., Heidmann, S., Reed, R., Myler, P. J., Goringer, H. U., and Stuart, K. D. (1998) *Mol Cell Biol* **18**, 6014-6022
214. Reichert, V. L., Le Hir, H., Jurica, M. S., and Moore, M. J. (2002) *Genes Dev* **16**, 2778-2791
215. Milligan, J. F., Groebe, D. R., Witherell, G. W., and Uhlenbeck, O. C. (1987) *Nucleic Acids Res* **15**, 8783-8798
216. Harris, M. E., and Hajduk, S. L. (1992) *Cell* **68**, 1091-1099
217. Sabatini, R., and Hajduk, S. L. (1995) *J Biol Chem* **270**, 7233-7240
218. Hema, M., and Kao, C. C. (2004) *J Virol* **78**, 1169-1180
219. Pai, R. D., Oppegard, L. M., and Connell, G. J. (2003) *RNA* **9**, 469-483
220. Oppegard, L. M., and Connell, G. J. (2002) *Int J Parasitol* **32**, 859-866
221. Panigrahi, A. K., Gygi, S. P., Ernst, N. L., Igo, R. P., Jr., Palazzo, S. S., Schnaufer, A., Weston, D. S., Carmean, N., Salavati, R., Aebersold, R., and Stuart, K. D. (2001) *Mol Cell Biol* **21**, 380-389
222. Huang, C. E., O'Hearn, S. F., and Sollner-Webb, B. (2002) *Mol Cell Biol* **22**, 3194-3203
223. Leung, S. S., and Koslowsky, D. J. (2001) *RNA* **7**, 1803-1816
224. Leung, S. S., and Koslowsky, D. J. (2001) *Nucleic Acids Res* **29**, 703-709
225. Oppegard, L. M., Hillestad, M., McCarthy, R. T., Pai, R. D., and Connell, G. J. (2003) *J Biol Chem* **278**, 51167-51175
226. Koslowsky, D. J., Reifur, L., Yu, L. E., and Chen, W. (2004) *RNA Biol* **1**, 28-34
227. Blum, B., Bakalara, N., and Simpson, L. (1990) *Cell* **60**, 189-198
228. Leung, S. S., and Koslowsky, D. J. (1999) *Nucleic Acids Res* **27**, 778-787
229. Salavati, R., Ernst, N. L., O'Rear, J., Gilliam, T., Tarun, S., Jr., and Stuart, K. (2006) *RNA* **12**, 819-831
230. Lawson, S. D., Igo, R. P., Jr., Salavati, R., and Stuart, K. D. (2001) *RNA* **7**, 1793-1802
231. Kapushoc, S. T., and Simpson, L. (1999) *RNA* **5**, 656-669
232. Kabb, A. L., Oppegard, L. M., McKenzie, B. A., and Connell, G. J. (2001) *Nucleic Acids Res* **29**, 2575-2580
233. Oppegard, L. M., Kabb, A. L., and Connell, G. J. (2000) *J Biol Chem* **275**, 33911-33919
234. Yu, L. E., and Koslowsky, D. J. (2006) *RNA* **12**, 1050-1060
235. Moore, M. J., and Sharp, P. A. (1992) *Science* **256**, 992-997
236. Nandakumar, J., and Shuman, S. (2004) *Mol Cell* **16**, 211-221
237. Arnott, S. (1970) *Prog Biophys Mol Biol* **21**, 265-319

238. Zuker, M. (2003) *Nucleic Acids Res* **31**, 3406-3415
239. Madison-Antenucci, S., and Hajduk, S. L. (2001) *Mol Cell* **7**, 879-886
240. Tian, B., Bevilacqua, P. C., Diegelman-Parente, A., and Mathews, M. B. (2004) *Nat Rev Mol Cell Biol* **5**, 1013-1023
241. Stefl, R., Skrisovska, L., and Allain, F. H. (2005) *EMBO Rep* **6**, 33-38
242. Wang, A. H., Fujii, S., van Boom, J. H., van der Marel, G. A., van Boeckel, S. A., and Rich, A. (1982) *Nature* **299**, 601-604
243. Messias, A. C., and Sattler, M. (2004) *Acc Chem Res* **37**, 279-287
244. Hudson, B. P., Martinez-Yamout, M. A., Dyson, H. J., and Wright, P. E. (2004) *Nat Struct Mol Biol* **11**, 257-264
245. O'Hearn, S. F., Huang, C. E., Hemann, M., Zhelonkina, A., and Sollner-Webb, B. (2003) *Mol Cell Biol* **23**, 7909-7919
246. Tarun, S. Z., Jr., Schnaufer, A., Ernst, N. L., Proff, R., Deng, J., Hol, W., and Stuart, K. (2008) *RNA* **14**, 347-358
247. Wirtz, E., Leal, S., Ochatt, C., and Cross, G. A. (1999) *Mol Biochem Parasitol* **99**, 89-101
248. Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Seraphin, B. (1999) *Nat Biotechnol* **17**, 1030-1032
249. Zeng, Y., and Cullen, B. R. (2005) *J Biol Chem* **280**, 27595-27603
250. Calin-Jageman, I., and Nicholson, A. W. (2003) *Nucleic Acids Res* **31**, 2381-2392
251. Reifur, L., and Koslowsky, D. J. (2008) *RNA* **14**, 2195-2211
252. Aphasizhev, R., Sbicego, S., Peris, M., Jang, S. H., Aphasizheva, I., Simpson, A. M., Rivlin, A., and Simpson, L. (2002) *Cell* **108**, 637-648
253. Fisk, J. C., Ammerman, M. L., Presnyak, V., and Read, L. K. (2008) *J Biol Chem* **283**, 23016-23025
254. Acestor, N., Panigrahi, A. K., Carnes, J., Zikova, A., and Stuart, K. D. (2009) *RNA* **15**, 277-286
255. Hashimi, H., Cicova, Z., Novotna, L., Wen, Y. Z., and Lukes, J. (2009) *RNA* **15**, 588-599
256. Wickstead, B., Ersfeld, K., and Gull, K. (2002) *Mol Biochem Parasitol* **125**, 211-216
257. Florens, L., Carozza, M. J., Swanson, S. K., Fournier, M., Coleman, M. K., Workman, J. L., and Washburn, M. P. (2006) *Methods* **40**, 303-311
258. Washburn, M. P., Wolters, D., and Yates, J. R., 3rd. (2001) *Nat Biotechnol* **19**, 242-247
259. Wolters, D. A., Washburn, M. P., and Yates, J. R., 3rd. (2001) *Anal Chem* **73**, 5683-5690
260. Wohlschlegel, J. A. (2009) Identification of SUMO-Conjugated Proteins and their SUMO Attachment Sites Using Proteomic Mass Spectrometry in *Methods in Molecular Biology*. pp 33-49
261. Eng, J. K., McCormack, A. L., and Yates, J. R. I. (1994) *Journal of the American Society for Mass Spectrometry* **5**, 976-989

262. Tabb, D. L., McDonald, W. H., and Yates, J. R., 3rd. (2002) *J Proteome Res* **1**, 21-26
263. Elias, J. E., and Gygi, S. P. (2007) *Nat Methods* **4**, 207-214
264. Robb, G. B., and Rana, T. M. (2007) *Mol Cell* **26**, 523-537
265. Margossian, S. P., Li, H., Zassenhaus, H. P., and Butow, R. A. (1996) *Cell* **84**, 199-209
266. Fairman, M. E., Maroney, P. A., Wang, W., Bowers, H. A., Gollnick, P., Nilsen, T. W., and Jankowsky, E. (2004) *Science* **304**, 730-734
267. Hermann, T., Schmid, B., Heumann, H., and Goringer, H. U. (1997) *Nucleic Acids Res* **25**, 2311-2318
268. Mingler, M. K., Hingst, A. M., Clement, S. L., Yu, L. E., Reifur, L., and Koslowsky, D. J. (2006) *Mol Biochem Parasitol* **150**, 37-45
269. Pusnik, M., Small, I., Read, L. K., Fabbro, T., and Schneider, A. (2007) *Mol Cell Biol* **27**, 6876-6888
270. Osato, D., Rogers, K., Guo, Q., Li, F., Richmond, G., Klug, F., and Simpson, L. (2009) *RNA* **15**, 1338-1344
271. Crick, F. (1970) *Nature* **227**, 561-563
272. Urlaub, H., Kuhn-Holsken, E., and Luhrmann, R. (2008) *Methods Mol Biol* **488**, 221-245
273. Wang, Z., Tollervey, J., Briese, M., Turner, D., and Ule, J. (2009) *Methods* **48**, 287-293
274. Ule, J., Jensen, K., Mele, A., and Darnell, R. B. (2005) *Methods* **37**, 376-386

VITA

Name: Alfredo J. Hernandez

Address: Department of Biochemistry and Biophysics,
Texas A&M University
2128 TAMU, College Station, TX 77843

Email Address: zerophobe@tamu.edu

Education: B.S., The University of Texas - Pan American, 2004

Ph.D., Texas A&M University, 2010

Publications

Sacharidou, A., Cifuentes-Rojas, C., Halbig, K., Hernandez, A., Dangott, L. J., De Nova-Ocampo, M., and Cruz-Reyes, J. (2006) *RNA* **12**, 1219-1228

Cifuentes-Rojas, C., Pavia, P., Hernandez, A., Osterwisch, D., Puerta, C., and Cruz-Reyes, J. (2007) *J Biol Chem* **282**, 4265-4276

Cruz-Reyes J., and Hernandez, A. (2008) in *RNA and DNA Editing* (Smith, H. C. ed.), John Wiley & Sons, Inc., New Jersey. pp 71–98.

Hernandez, A., Panigrahi, A., Cifuentes-Rojas, C., Sacharidou, A., Stuart, K., and Cruz-Reyes, J. (2008) *J Mol Biol* **381**, 35-48

Hernandez, A.*, Madina, B. R.*, Ro, K., Wohlschlegel, J. A., Willard, B., Kinter, M. T., and Cruz-Reyes, J. (2010) *J Biol Chem* **285**, 1220-1228
**Equal contribution.*