



THESIS

3  
2009

**LIBRARY**  
**Michigan State**  
**University**

This is to certify that the  
dissertation entitled

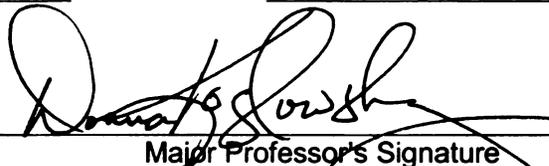
**STRUCTURAL AND THERMODYNAMIC STUDIES OF THE  
ATPASE SUBUNIT 6 MRNA/GRNA COMPLEX IN  
*TRYPANOSOMA BRUCEI***

presented by

**LARISSA REIFUR**

has been accepted towards fulfillment  
of the requirements for the

PhD degree in Comparative Medicine and  
Integrative Biology

  
Major Professor's Signature  
12/8/2008  
Date

**PLACE IN RETURN BOX** to remove this checkout from your record.  
**TO AVOID FINES** return on or before date due.  
**MAY BE RECALLED** with earlier due date if requested.

DATE DUE	DATE DUE	DATE DUE

**STRUCTURAL AND THERMODYNAMIC STUDIES OF  
THE ATPASE SUBUNIT 6 MRNA/GRNA COMPLEX IN  
*TRYPANOSOMA BRUCEI***

By

Larissa Reifur

A DISSERTATION

Submitted to  
Michigan State University  
in partial fulfillment of the requirements  
for the degree of

DOCTOR OF PHILOSOPHY

Comparative Medicine and Integrative Biology

2008

## ABSTRACT

### STRUCTURAL AND THERMODYNAMIC STUDIES OF THE ATPase SUBUNIT 6 mRNA/gRNA COMPLEX IN *TRYPANOSOMA BRUCEI*

By

Larissa Reifur

Mitochondrial mRNA editing in *Trypanosoma brucei* is a unique post-transcriptional process that controls mRNA maturation and gene expression. Editing requires diverse enzyme activities present in the editosome, a multi-protein complex that utilizes the information encoded in small guide RNAs (gRNA) to insert and delete uridylates within specific mRNA sites. The complete mechanism for mRNA/gRNA complex formation and editing site recognition by the editosome is unknown. Since RNA structure is essential for numerous enzymatic processes, we hypothesize that distinct mRNA/gRNA complexes share common structural features important for editing. Consensus sequences have not been found, but crosslinking studies have shown that different mRNA/gRNAs can fold into a comparable three-helical configuration. Through solution structure probing, this architectural organization proved to exist in the apocytochrome b (CYb) mRNA/gRNA complex, supporting our hypothesis.

The present dissertation was divided according to three projects, with the first project comprising most of the work. **1.** Structural modeling of *T. brucei* ATPase subunit 6 mRNA (A6) bound to its cognate guide RNA (gA6-14) using site specific crosslinking and solution structure probing. Here we determined the A6/gA6-14 secondary structure and compared with the CYb/gCYb-558 structure. **2.** The impact of mRNA structure on guide RNA targeting in kinetoplastid RNA editing. Here we examined the accessibility of several gRNAs to different mRNAs using computer algorithms for RNA folding,

RNase H experiments, electrophoretic mobility shift assays, solution structure probing, surface plasmon resonance, and gRNA directed cleavage assays. 3. The influence of the gRNA U-tail on the thermodynamics of A6/gA6-14 binding. We utilized isothermal titration calorimetry and electrophoretic mobility shift assays and determined the binding affinity, the changes in enthalpy, entropy, and free energy that are coupled to binding.

The work presented in this dissertation describes how the gRNA interacts with the mRNA, the influence of their structures in the binding event, and the structure adopted by the mRNA/gRNA complex. Together, these data contribute to the understanding of common and specific RNA elements necessary for assembly of mRNA/gRNA complexes. While this study brings important information to the mechanisms of mRNA editing in kinetoplastid parasites it also contributes to the better understanding of RNA-RNA interaction in general.

**I dedicate this thesis to my parents and to my sister for all the support and care**

## ACKNOWLEDGMENTS

I would like to first thank my family. If acquiring the Veterinary degree, the master's and the Ph.D. degree means going up a few steps on my life-ladder, then my parents, Nilo and Irene, are the ones holding the ladder, the most important persons in my life. Equally, I thank my sister Melissa, a fabulous, dedicated, and compassionate friend that helped me immensely throughout the program.

I gratefully acknowledge Donna Koslowsky, my advisor, for her supervision, advice, and guidance that allowed me to succeed as a student and to grow as a researcher and a scientist.

I would like to thank the guidance committee members, Drs. John Fyfe, Charles Hoogstraten, Vilma Yuzbasiyan-Gurkan, and John Wang for their critiques and recommendations. Part of my project I spent using Dr. Hoogstraten's ITC equipment and I thank him for the training, for letting me use it, and for all he taught me on RNA and thermodynamics, including the long last-minute e-mails.

All these years working in the Koslowsky laboratory would not be the same without the former and current students. Sandra Clement was a great friend and also a mentor. With her happiness, Melissa Mingler brought much joy to the lab, even when something went wrong, she would smile and continue. Laura E. Yu was the pivotal person; she taught me many techniques and helped me enormously throughout the 5 years we spent together. Laura and I were alike in many ways, so we became the right hand of each other (as she wrote in her thesis). I thank them all and also all

undergraduates that indirectly helped with the project, especially Tanojo Mustika, Andrea Hingst, Meghan Vanslager, and Remy Brim.

I have learned much participating of the MSU RNA Journal Club and I thank everyone from Dr. Patterson's, Dr. Koslowsky's, Dr. Hoogstraten's, and Dr. Wang's laboratories, for their listening and helpful ideas.

I would like to thank Dr. Linda Mansfield for being my mentor on MMG890, special problems in microbiology and continuing being my mentor until today. In addition to learning about *Leishmania* spp., I had a great experience working as a teaching assistant (TA) for the MMG569 Veterinary Parasitology wet lab.

I thank the Comparative Medicine and Integrative Biology Program, the College of Veterinary Medicine, the MMG department, the Graduate School, and the Office for International Students and Scholars at MSU, for the support in all means, including financial. I would like to thank the CMIB Director, Dr. Yuzbasiyan-Gurkan, for her compassion, support, and guidance. I thank Victoria Hoelzer-Maddox for her help within the CMIB.

Many thanks go to the extraordinary people who took the Biology of Parasitism course with me in Woods Hole: Angie, Arnaud, Beth, Bruna, Isabel, Jo, Katy, Mai, Nick, Najju, Paul, Tiff, the course organizers Drs. Patricia Johnson, Boris Striepen, Judi Allen, Debbie Smith, Meg Phillips, the CAs and TAs, and the speakers that made BoP 2006 an outstanding course. We learned and grew much together. BoP tremendously improved my professional and personal skills.

Finally, I would like to thank Drs. Fabiano M. Ferreira and Vanete Thomaz-Soccol, for their encouragement.

## TABLE OF CONTENTS

LIST OF TABLES .....	XI
LIST OF FIGURES .....	XII
LIST OF EQUATIONS .....	XIV
LIST OF EQUATIONS .....	XIV
KEY TO ABBREVIATIONS .....	XV
<b>CHAPTER 1 .....</b>	<b>1</b>
<b>INTRODUCTION TO KINETOPLASTID PARASITES, TRYPANOSOMA BRUCEI BIOLOGY AND RNA EDITING .....</b>	<b>1</b>
TRYPANOSOMATIDS .....	2
TRYPANOSOMA BRUCEI.....	11
<i>Procyclic form (PF)</i> .....	11
<i>Epimastigote</i> .....	11
<i>Metacyclic</i> .....	12
<i>Bloodstream form (BF)</i> .....	12
<i>Cell and metabolism</i> .....	14
<i>Flagellar pocket</i> .....	14
<i>Flagellum</i> .....	14
<i>Glycosome</i> .....	15
<i>Mitochondrion (mt)</i> .....	18
<i>Kinetoplast (kDNA)</i> .....	22
RNA EDITING .....	26
<i>Control of mitochondrial biogenesis</i> .....	26
<i>Mitochondrial-encoded Adenosine triphosphatase subunit 6 (A6)</i> .....	28
<i>gRNAs</i> .....	33
<i>Editing machinery (editosome)</i> .....	33
<i>Accessory proteins</i> .....	38
<i>RNA editing mechanism</i> .....	39
OTHER RELEVANT RNA SYSTEMS .....	42

<i>Antisense RNA</i> .....	43
<i>Micro-RNAs</i> .....	46
<i>Other RNA-RNA interactions including RNA editing in other systems</i> .....	48
CONCLUDING REMARKS.....	51
FOCUS OF THIS DISSERTATION .....	54
REFERENCES .....	57
<b>CHAPTER 2 .....</b>	<b>75</b>
<b><i>TRYPANOSOMA BRUCEI</i> ATPASE SUBUNIT 6 MRNA BOUND TO GA6-14 FORMS A CONSERVED THREE-HELICAL STRUCTURE.....</b>	<b>75</b>
ABSTRACT .....	76
INTRODUCTION.....	76
RESULTS.....	80
<i>Description of RNAs</i> .....	80
<i>gRNA U-tail interactions with mRNA</i> .....	81
<i>Solution Structure Probing</i> .....	91
DISCUSSION.....	106
MATERIALS AND METHODS.....	111
<i>Oligodeoxynucleotides (ODN)</i> .....	111
<i>Oligoribonucleotides</i> .....	111
<i>DNA Templates and RNA synthesis and modification</i> .....	111
<i>RNA Crosslinking</i> .....	113
<i>Primer Extension Analysis</i> .....	113
<i>RNase H assays</i> .....	114
<i>Secondary Structure Prediction</i> .....	114
<i>Solution Structure Probing</i> .....	114
ACKNOWLEDGMENTS .....	116
REFERENCES .....	117
<b>CHAPTER 3 .....</b>	<b>124</b>
<b>THE IMPACT OF MRNA STRUCTURE ON GUIDE RNA TARGETING IN KINETOPLASTID RNA EDITING .....</b>	<b>124</b>
ABSTRACT .....	125
INTRODUCTION.....	125
RESULTS.....	128
<i>Description of mRNAs and analyses of their predicted secondary structures</i> .....	128

<i>Determination of ABS accessibility</i> .....	130
<i>Analysis of binding affinities</i> .....	132
<i>mRNA/gRNA rate constants</i> .....	136
<i>Ability of each mRNA/gRNA pair to be recognized and cleaved by the editosome</i> .....	138
DISCUSSION.....	143
MATERIALS AND METHODS.....	148
<i>Oligonucleotides</i> .....	148
<i>Templates for RNA transcription</i> .....	149
<i>RNA transcription and radioactive 5'-end-labeling</i> .....	149
<i>Secondary Structure Prediction</i> .....	150
<i>ODN directed RNase H assays</i> .....	150
<i>Determination of Binding Affinity</i> .....	151
<i>Determination of rate constants</i> .....	152
<i>gRNA Directed Cleavage Assays</i> .....	154
<i>Solution Structure Probing</i> .....	155
ACKNOWLEDGEMENTS.....	156
REFERENCES.....	157
<b>CHAPTER 4</b> .....	<b>164</b>
<b>U-TAIL INFLUENCE ON THE BINDING THERMODYNAMICS OF ATPASE SUBUNIT 6 MRNA/GRNA IN <i>TRYPANOSOMA BRUCEI</i></b> .....	<b>164</b>
ABSTRACT.....	165
INTRODUCTION.....	165
RESULTS.....	168
<i>Description of the RNAs</i> .....	168
<i>Individual RNA structures and Computer simulations</i> .....	169
<i>mRNA/gRNA structure</i> .....	172
<i>Determination of binding affinity</i> .....	173
<i>Thermodynamics of mRNA/gRNA binding</i> .....	175
DISCUSSION.....	183
MATERIALS AND METHODS.....	188
<i>Oligodeoxynucleotides (ODN)</i> .....	188
<i>RNA Synthesis and Labeling</i> .....	188
<i>Secondary Structure Prediction</i> .....	189
<i>Isothermal Titration Calorimetry (ITC)</i> .....	189
<i>Electrophoretic Mobility Shift Assays (EMSA or gel shifts)</i> .....	192
ACKNOWLEDGEMENTS.....	193
REFERENCES.....	194

<b>CHAPTER 5 .....</b>	<b>198</b>
<b>CONCLUSIONS AND FUTURE DIRECTIONS.....</b>	<b>198</b>
<b>INTRODUCTION .....</b>	<b>199</b>
<b>SUMMARY OF RESULTS.....</b>	<b>200</b>
<i>Chapter 2 - A6 mRNA/gRNA secondary structure during editing .....</i>	<i>200</i>
<i>Chapter 3 - mRNA structure and gRNA targeting .....</i>	<i>201</i>
<i>Chapter 4 - mRNA/gRNA binding thermodynamics.....</i>	<i>202</i>
<b>DISCUSSION, CONCLUSIONS AND FUTURE DIRECTIONS .....</b>	<b>202</b>
<b>REFERENCES .....</b>	<b>207</b>

## LIST OF TABLES

TABLE 1.1 PREVALENCE, INCIDENCE AND BURDEN (DALY) OF IMPORTANT NEGLECTED DISEASES IN THE WORLD IN 2002.....	4
TABLE 1.2 MITOCHONDRIAL EDITING MODIFICATION FOR EACH OF THE <i>T. BRUCEI</i> MRNAs. ....	28
TABLE 1.3 LIST OF <i>T. BRUCEI</i> EDITOSOME PROTEINS ISOLATED FROM U-INSERTION AND U-DELETION EDITOSOMES, THEIR FUNCTION AND FUNCTIONAL DOMAINS. ....	35
TABLE 2.1 LIST OF OLIGODEOXYNUCLEOTIDES .....	111
TABLE 3.1 MRNA/GRNA APPARENT BINDING AFFINITIES. THE $K_D$ WAS ALSO CALCULATED WHEN THE GRNAS HAD THE U-TAIL DELETED (NO U-TAIL). ....	133
TABLE 3.2 OLIGODEOXYRIBONUCLEOTIDES.....	148
TABLE 4.1 PREDICTED CHANGES IN FREE ENERGY ( $\Delta G$ ) AND MELTING TEMPERATURE ( $T_M$ ) FOR THE A6 MRNA/GRNA COMPLEXES USED IN THIS STUDY. * .....	173
TABLE 4.2 THERMODYNAMIC PARAMETERS FROM ISOTHERMAL TITRATION CALORIMETRY. ....	180
TABLE 4.3 LIST OF OLIGODEOXYNUCLEOTIDES .....	188

## LIST OF FIGURES

FIGURE 1.1 <i>T. BRUCEI</i> LIFE CYCLE THROUGH THE MAMMALIAN (37°C) AND INSECT (27°C) HOSTS. ....	13
FIGURE 1.2 SCHEMATICS OF ENERGY METABOLISM IN THE BLOODSTREAM FORM <i>T. BRUCEI</i> . ....	16
FIGURE 1.3 SCHEMATICS OF ENERGY METABOLISM IN THE INSECT STAGE, OR PROCYCLIC FORM <i>T. BRUCEI</i> . .	18
FIGURE 1.4 SCHEMATICS OF THE MITOCHONDRIAL RESPIRATION COMPLEXES IN <i>T. BRUCEI</i> . ....	21
FIGURE 1.5 SCHEMATICS OF <i>T. BRUCEI</i> MAXICIRCLE, MINICIRCLE, GRNA STRUCTURE, EDITING PROGRESSION, AND EDITING STEPS. ....	25
FIGURE 1.6 DNA AND FULLY EDITED mRNA SEQUENCES OF THE ATPASE SUBUNIT 6 FROM <i>T. BRUCEI</i> AND <i>L.</i> <i>TARENTOLAE</i> . ....	30
FIGURE 1.7 SECONDARY STRUCTURE MODEL FOR THE mRNA/GRNA COMPLEX. ....	56
FIGURE 2.1 A6 MRNAs AND GRNAs. ....	81
FIGURE 2.2 A6U AND A6P3 CROSSLINKS WITH 5' <sup>32</sup> P-LABELED GA6-14-U5, (U5) OR GA6-14-U10, (U10). 83	83
FIGURE 2.3 MAPPED CROSSLINK SITES WITHIN A6U AND A6P3 CROSSLINKED TO GA6-14-U5 OR GA6-14- U10 (GRNA SEQUENCE IS NOT SHOWN). ....	85
FIGURE 2.4 RNASE H ASSAYS CONFIRMING CROSSLINK SITES FOR B1 AND B2 POPULATIONS OF A6U AND A6P3 CROSSLINKED TO GA6-14-U5 (U5) OR GA6-14-U10 (U10). ....	87
FIGURE 2.5 SOLUTION STRUCTURE PROBING. 5'-END LABELED A6U AND A6P3 ALONE OR HYBRIDIZED TO GA6-14 (+14), GA6-14sU (+sU), AND GA6-14-C6 (+C6). ....	96
FIGURE 2.6 SECONDARY STRUCTURE MODELS FOR THE INDIVIDUAL A6 RNAs AND mRNA/GRNAs DURING DISTINCT STAGES OF EDITING. ....	99
FIGURE 3.1 MRNAs ALIGNED WITH GRNAs AND ODNS. ....	129

FIGURE 3.2 PREDICTED SECONDARY STRUCTURES FOR A6U, CYbU, ND7UHR3, ND7-550.....	130
FIGURE 3.3 ODN-DIRECTED ACCESSIBILITY ASSAYS.....	132
FIGURE 3.4 MRNA/GRNA BINDING AFFINITY BY EMSA.....	135
FIGURE 3.5 MRNA/GRNA SENSORGRAMS AND RATE CONSTANTS BY SURFACE PLASMON RESONANCE. ....	138
FIGURE 3.6 IN VITRO GRNA-DIRECTED CLEAVAGE ASSAY.....	140
FIGURE 3.7 PREDICTED SECONDARY STRUCTURES FOR THE A6P1/GA6-14 AND ND7-550/GND7-550 COMPLEXES. (*) SITES WHERE WE OBSERVED GRNA-DIRECTED CLEAVAGES (C1 – C5). ....	142
FIGURE 3.8 SOLUTION STRUCTURE PROBING OF ND7-550/GND7-550.....	143
FIGURE 4.1 MRNA AND GRNA SEQUENCES.....	169
FIGURE 4.2 SECONDARY STRUCTURES OF THE MRNA, GRNAS AND COMPLEXES.....	170
FIGURE 4.3 MRNA/GRNA COMPLEX FORMATION BY EMSA.....	175
FIGURE 4.4 ITC DATA FOR A6P1 ANNEALING WITH GA6-14sU, GA6-14, AND GA6-14-C6.....	181
FIGURE 4.5 SCHEMATICS OF STRUCTURAL REARRANGEMENTS PREDICTED TO HAPPEN IN THE THERMODYNAMICS OF THREE DISTINCT MRNA-GRNA BINDING EVENTS.....	187

## ***LIST OF EQUATIONS***

EQUATION 3.1 DISSOCIATION EQUILIBRIUM CONSTANT .....	151
EQUATION 3.2 DISSOCIATION EQUILIBRIUM CONSTANT FOR THE MRNA/GRNA COMPLEX .....	151
EQUATION 3.3 ASSOCIATION RATE CONSTANT .....	153
EQUATION 3.4 DISSOCIATION RATE CONSTANT .....	153
EQUATION 3.5 DISSOCIATION EQUILIBRIUM CONSTANT .....	154
EQUATION 4.1 ASSOCIATION AFFINITY CONSTANT .....	191
EQUATION 4.2 CONCENTRATION OF LIGAND X .....	191
EQUATION 4.3 FRACTION OF SITES OCCUPIED BY THE LIGAND X .....	191
EQUATION 4.4 TOTAL HEAT .....	191
EQUATION 4.5 TOTAL HEAT FOR AN INJECTION .....	191
EQUATION 4.6 HEAT RELEASED FROM A INJECTION .....	192

## KEY TO ABBREVIATIONS

A	adenosine
AAT	animal African trypanosomosis
ABS	anchor binding site (mRNA domain)
APS	azidophenacyl bromide
A6	ATPase subunit 6 mRNA
A6U	truncated ATPase subunit 6 mRNA Unedited
A6P3	truncated ATPase subunit 6 mRNA Partially Edited through Site 3
BS	blood stream form
C	cytosine
CYbU	truncated CYb mRNA Unedited
CYbPES3	truncated CYb mRNA Partially Edited through Site 3
DEPC	diethyl pyrocarbonate
G	guanosine
Ds	double stranded
ES	editing site
gA6-14	5'-most initiating A6 gRNA
gA6-14-C6	gA6-14 with mutated U-tail (6 U-to-C mutations)
gA6-14sU	gA6-14 without the U-tail
gCYb-558	initiating CYb gRNA
gRNA	guide RNA
HAT	human African trypanosomosis
ITC	isothermal titration calorimetry
K <sub>D</sub>	dissociation constant
k <sub>obs</sub>	observed rate constant
k <sub>off</sub>	dissociation rate constant
k <sub>on</sub>	association rate constant
Nt	nucleotide (s)
ODN	oligodeoxynucleotide
PF	procyclic form
RNase	ribonuclease
RT	reverse transcriptase
RU	resonance unit
SPR	surface plasmon resonance
Ss	single stranded
T1	RNase T1
T2	RNase T2
U	uridine
U-tail	poly(U) tail
UV	ultra-violet light
V1	RNase V1
VSG	variant surface glycoprotein

## **Chapter 1**

# **INTRODUCTION TO KINETOPLASTID PARASITES, TRYPANOSOMA BRUCEI BIOLOGY AND RNA EDITING**

## ***Trypanosomatids***

The order Kinetoplastida is formed of single-celled flagellated protozoans found free-living or parasitizing insects, plants, and vertebrates (Simpson et al. 2006).

Kinetoplastids are among the most ancient eukaryotes with a mitochondrion. Their single mitochondrion contains a structurally complex DNA called kinetoplast or kDNA (Opperdoes and Michels 2007). Kinetoplastids are divided into Protokinetoplastina (including *Ichthyobodo*, an ectoparasite of fish, and *Perkinsiella*, a small endosymbiont of certain amoebae) and Metakinetoplastina [divided in the groups neobodonida, parabodonida (including the fish-infecting *Trypanoplasma borelli*), eubodonida (including the free-living *Bodo saltans*), and trypanosomatida] (Simpson et al. 2006).

Trypanosomatids are parasitic and divided in several clades such as *T. brucei*, *T. cruzi*, rodent, avian, and aquatic clades. Also within trypanosomatids are genera *Leishmania*, *Phytomonas*, and other medically unimportant trypanosomatids. *Phytomonas* spp. are plant parasites transmitted by the saliva of phytophagous hemipterous insects. They can destroy entire crops of high economic interest like coffee, coconut and oil palm (Camargo 1999). Of medical importance are the insect-vectored genera *Trypanosoma* and *Leishmania*, responsible for causing some of the most neglected zoonotic diseases in the world. *Trypanosoma cruzi* is transmitted via feces of triatomine bugs and causes Chagas disease in approximately 10 million people, killing around 13 thousand per year in the Americas (Barrett et al. 2003). In the United States, 100 thousand people have been diagnosed with Chagas disease. In addition, *T. cruzi*-infected triatomine vectors and animals have also been identified in the US, however no autochthonous transmission has been confirmed (Bern et al. 2007). *Leishmania* spp. are

transmitted by phlebotomine sandflies and cause a range of cutaneous, mucocutaneous, and visceral diseases. Currently, approximately 350 million people in 88 tropic and subtropic countries, including South and Central America, the Mediterranean, and Asia are at risk. In the United States, there was an outbreak in 2000 where *L. infantum* was isolated from foxhounds in 18 states and in Canada. The parasite was mostly infecting foxhounds from hunt clubs (42% were positive for antibody to leishmania) and only 2% of other dogs around the USA (Grosjean et al. 2003; Rosypal 2005). This raised concerns that the parasite could take its way into the human population since dogs are reservoirs of *L. infantum*. However, no autochthonous human leishmaniasis has been detected (Rosypal 2005). *T. equiperdum* and *T. evansi* are responsible for dourine and surra, respectively, economically significant diseases infecting horses, camels, and water buffaloes, transmitted by bloodsucking insects or coitus (Brun et al. 1998; Naessens 2006). *T. evansi* is widely spread in Africa, Asia, and South America and causes disease mainly in camels, but can also parasitize horses, humans and other mammals. *T. congolense* and *T. vivax* are important tsetse-transmitted-parasites afflicting African livestock. *T. vivax* is also in South America, where it is transmitted by biting flies.

### ***Trypanosoma brucei***

In Africa, subspecies of *T. brucei* are responsible for causing devastating and neglected diseases in humans and animals. *T. brucei gambiense* causes a chronic form of human African trypanosomiasis (HAT), or sleeping sickness, in West and Central Africa, whereas *T. brucei rhodesiense* causes an acute form of HAT in East and Southern Africa (Barrett et al. 2003). HAT afflicts 200 thousand people, with 46,000 new cases per year (Table 1.1). The World Health Organization (WHO, <http://www.who.int/en/>) alerts that

these numbers are underestimated and 0.5 million people may be infected with 300,000 new cases occurring yearly (WHO 2001). These numbers appear small in contrast to other neglected diseases but HAT has high Disability Adjusted Life Year (DALY), relatively high control and treatment costs (US \$10 per DALY, \$120 per patient) and high social and economic impact that are not considered in the treatment costs and DALY measurements (Cattand et al. 2001; WHO 2001). DALY is a standard unit of health measurement used to estimate disease burden or the impact of a health problem in an area. Mathers and collaborators explain how DALY is measured by summing the premature mortality (years of life lost, YLL) and disability (years of life lived with a disability, YLD) (Mathers et al. 2007). Briefly, YLL takes in consideration the number of deaths multiplied by the life expectancy as a function of each age. YLD multiplies the number of cases at each age and time period by the average duration of the disease for each age and by a weight factor that reflects the severity of the disease. Despite the uncertainties, the numbers presented on table 1.1 give comparative information on mortality and disease burden in respect to HAT and other tropical diseases.

Table 1.1 Prevalence, incidence and burden (DALY) of important neglected diseases in the world in 2002.

Disease	Prevalence*	Incidence*	Deaths*	YLL*	YLD*	DALY*
Malaria	4,406	408,250	1,272	41,507	4,979	46,486
Leishmaniasis			51	1,569	521	2,090
Visceral	1,508	534				
Cutaneous	2,157	1,157				
HAT	200	46	48	1,429	96	1,525
Chagas	10,137	217	14	185	481	667

\*Multiply all numbers by  $10^3$ .

One lost DALY: one lost year of "healthy" life (through either death or illness/disability).

Reference: Mathers et al. (2007).

Fevre et al. suspect that in some malaria-endemic-places in Africa, malaria is diagnosed solely based on clinical signs, suggesting that other fever-presenting diseases

(including early stage of HAT) may be diagnosed as malaria, over-estimating malaria by up to 57% (Fevre et al. 2008).

Both human African trypanosomes are morphologically similar to the animal African trypanosome *T. b. brucei* (here referred as *T. brucei*) that causes nagana in livestock, restricting agricultural development and increasing poverty. *T. brucei* is not infectious to humans because it does not encode the SRA gene (serum resistance associated protein) present in *T. b. rhodesiense*. SRA confers resistance to lysis by apolipoprotein L1, a human component of the HDL fraction termed trypanosome lytic factor (Oli et al. 2006). On other hand, *T. b. rhodesiense* survives in the cattle blood and also other domestic animals, which serve as reservoirs together with wild animals (Naessens 2006).

All subspecies of *T. brucei* that infect humans or animals are able to cause re-emergent epidemics because the existence of vectors and reservoirs. The natural vector is the tsetse fly (genus *Glossina*, order Diptera), present within the tsetse fly belt of approximately ten million square kilometers in equatorial Africa, between 14° North and 29° South (Njiokou et al. 2004). Nonvector transmission occasionally occurs by blood transfusion, contaminated needles, and via placenta (Barrett et al. 2003). Controlling the population of flies has been the most effective form of reducing the incidence of African trypanosomosis, since there are no vaccines and drugs are inefficient (Jannin and Cattand 2004). Wild and domestic animals are the main reservoir hosts of *T. brucei* and *T. b. gambiense* whereas humans and cattle carriers are important reservoirs of *T. b. rhodesiense* (WHO 2001; Gibson 2007).

## **Human African trypanosomiasis or sleeping sickness**

HAT is caused by distinct *Trypanosoma brucei* subspecies (*T. brucei gambiense* and *T. brucei rhodesiense*) and is transmitted by the tsetse fly. HAT starts with a local inflammatory reaction, swelling of the skin, and enlargement of the draining lymph node. This initial reaction is due to the tsetse bite and the trypanosomes that are deposited in the skin. The skin lesion (chancre) heals after 3 to 4 weeks while the parasites migrate to live and multiply in lymphatic and blood vessels. The pathophysiology of HAT is mainly due to an inflammatory response and is divided into stage 1, or haemolympathic, and stage 2, or meningoencephalopathic (Sternberg 2004). The rate of progression through stages 1 and 2 varies depending on the parasite. *T. b. gambiense* infection is chronic, the haemolympathic stage 1 lasts for 6 to 12 months and is characterized by periodical outbreaks due to cyclic antigenic variation. The immune complexes formed after parasite lysis lead to perivascular inflammatory signs like oedema. Other clinical signs are intermittent fever, headache, malaise, and generalized lymphadenopathy, especially at the neck (Winterbottom sign). The meningoencephalopathic stage 2 in the gambiense infection involves parasite invasion of internal organs and the CNS, lasting several months or years, and ending in death. Some clinical signs include weight loss, endocrine abnormalities, tremor, paraesthesia, increased sensitivity to pain, gait disorders, speech difficulties, personality changes, and reversal of the diurnal wake/sleep rhythm (sleeping sickness) among many other neurological signs. Immunosuppression, high nitric oxide and IL-10 levels in the cerebrospinal fluid are found in this stage. The CNS progressively deteriorates leading to stupor, coma, and death, accompanied by malnutrition and concomitant infections. Leukoencephalitis with autoimmune demyelination are the main histopathological findings; yet, trypanosomes are rarely seen

in the CNS (Barrett et al. 2003). The disease caused by *T. b. rhodesiense* evolves faster through the 2 stages, with only 1 to 3 weeks of incubation. Starting with a more severe reaction at the bite-site, high fever and signs of multiple organ failure, especially heart collapse, which usually proceeds to death within a few weeks or months (Barrett et al. 2003).

### **Animal African trypanosomosis (AAT) or nagana**

AAT is caused *Trypanosoma* species and transmitted by the tsetse fly. The pathophysiology is similar to what was described for HAT above. The initial lesion at the bite-site with enlarged lymph nodes is followed by the systemic infection and generalized inflammatory reaction. Hepatosplenomegaly, leucopenia, anemia, and loss of weight are typical findings. Animals with chronic infection become lethargic, lose appetite and die usually of congestive heart failure (Naessens 2006). *T. congolense* and *vivax* remain in the vessels and thus do not cause CNS disease. *T. brucei* spp. and *T. evansi* (also *T. b. rhodesiense*) invade other tissues, including the CNS, causing pathological findings similar to HAT meningoencephalopathic stage 2 (Naessens 2006).

### **Trypanotolerance**

The indigenous animals (e.g. N'Dama) of taurine (*Bos taurus*) origin developed tolerance to African trypanosomes through natural selection over several millennia. Although N'Dama can get infected, it acquired genetic capacity to better control parasitemia and anemia, allowing it to be able to survive and be productive in endemic regions. In contrast, the newly introduced (700 A.D.) zebu (e.g. Boran) (*Bos indicus*) in Africa is susceptible to infection and can only be raised in Africa with trypanocidal drugs. The precise mechanisms leading to trypanotolerance are unknown, but the control of both

anemia and parasitemia are independent of T cells and antibody (O'Gorman et al. 2006). The control of anemia is mediated by cells from the hemopoietic system, while control of parasitemia was not dependent on the same system and likely to be innate (Naessens 2006). The parasitemia control in trypanotolerant N'Dama may be due to earlier development of T<sub>H1</sub> response (O'Gorman et al. 2006).

### **Diagnosis of African trypanosomosis**

Due to low parasite loads, diagnosis is based only on clinical signs or in combination with microscopy for detection of parasite in blood, chancre or lymph node (Buscher 2002). In humans, card agglutination test for trypanosomosis (CATT) for detection of antibody in serum but further parasitological tests are required to confirm the diagnosis. Lumbar puncture to collect CSF (for parasitological examination, white blood cell count, protein concentration, and antibody detection) is essential for diagnosis of stage 2 (Chappuis et al. 2005). Other techniques like ELISA and IFA, in vitro culture and animal inoculation, detection of parasite DNA through PCR and PCR-based techniques have been developed with success but are more expensive (Deborggraeve et al. 2006).

### **Treatment**

The available drugs to treat African trypanosomosis are suramin, pentamidine, melarsoprol and eflornithine.

Suramin (or polysulphonated naphthyl urea) was released in 1920 to treat stage 1 patients. It is strongly negatively charged at physiological pH, thus binds and inhibits a multitude of enzymes. Unfortunately, suramin was not very effective against *T. b. rhodesiense*, *T. vivax*, and *T. congolense* and resistance of *T. evansi* to suramin was high (Delespaux and de Koning 2007).

Pentamidine replaced suramin and has been used since 1940. Resistance has not been demonstrated, only treatment failures due to use in misdiagnosed late stage patients. Drug uptake by the parasite is high and the mode of action is probably due to its electrostatic interaction with nucleic acids, since it is a di-cation (Docampo and Moreno 2003).

Melarsoprol is a trivalent arsenic compound used since 1947. The mode of action is unknown but glycerol-3-phosphate dehydrogenase is a potential target since it strongly binds to a melarsoprol analog (Nok 2003). It is lipophilic and diffuses through the blood brain barrier (Delespaux and de Koning 2007) thus, it is the standard drug used to treat patients in the meningoencephalopathic stage. However, the intravenous administration (dissolved in propylene glycol) is extremely painful, causes post-injection thrombophlebitis, and encephalopathy in 19% of patients. 2 to 12% of the patients die after receiving the drug (Barrett et al. 2003; Nok 2003).

Eflornithine ( $\alpha$ -difluoromethylornithine, DFMO) released in 1990, is the newest drug on the market. It is an analogue of ornithine, thus it inhibits ornithine decarboxylase, an enzyme used for the formation of polyamines and fundamental for cell proliferation (Delespaux and de Koning 2007). DFMO inhibits trypanosome division and also their capability of changing the variant surface glycoprotein, which makes them vulnerable targets for the host immune system. Eflornithine is an alternative to melarsoprol because it crosses the BBB and has fewer side effects but is only effective against *T. b. gambiense*. The main limiting factor in its use is the difficulty to administer because treatment requires a large dose administered in four daily intravenous infusions for 7-14 days (Barrett et al. 2003; Docampo and Moreno 2003).

All the above trypanocidal drugs induce unacceptably high levels of toxicity and resistance (Barrett et al. 2003). Thus, there is urgent need for new drugs and drug targets.

### **Vaccine**

No effective vaccine exists, and development has been difficult due to the ability of the parasite to evade the host immune response via antigenic variation.

### **Antigenic variation**

African trypanosomes live freely in the blood and can be cleared by antibodies.

However, their evasion mechanism, or antigenic variation, is potent and involves cyclical changes of the surface antigens. This antigenic coat is made of dense  $10^8$  variant surface glycoprotein (VSG) molecules attached to the cell membrane by a glycosylphosphatidylinositol (GPI) anchor (Sternberg 2004). VSG protects against complement-mediated lysis but is highly immunogenic (elicits T-cell dependent and independent B-cell responses) (Barrett et al. 2003). There are more than a thousand genes encoding VSG variants; however, only one type is expressed at a time, with a new VSG gene replacing the former active VSG gene at the transcriptionally active telomeric site. This mechanism allows the parasite to sequentially express antigenically distinct glycoproteins whenever specific antibodies are made and start lysing trypanosomes (Zambrano-Villa et al. 2002). In response, host polyclonal B-cell activation with generation of auto-antibodies and immune complex in addition to raised IgM and IgG are characteristics of HAT (Sternberg 2004).

### **Potential drug targets**

Although trypanosomes are eukaryotic cells, some metabolic pathways, membrane architecture, and organelles are distinct from mammalian cells. For instance,

trypanosome pentose phosphate pathway is related to cyanobacterial isoforms, sterol metabolism is similar to fungi, and both differ from that in mammalian cell. The mitochondrion of these parasites is a potential target because of its kinetoplast and RNA editing mechanism that are unique to trypanosomes (Schnauffer et al. 2001; Schnauffer et al. 2002; Stuart et al. 2002; Barrett et al. 2003). Glycosomes (described below) are important drug targets because are essential for *T. brucei* survival and are nonexistent in mammalian cells. Glycosomes contain glycolytic enzymes that differ from the mammalian enzymes in structure, organization, and regulation (Hellemond et al. 2005).

### ***Trypanosoma brucei***

As a unicellular eukaryotic organism, *T. brucei* undergoes extreme morphological and physiological changes to adapt to different environments as it cycles between the mammalian host and the insect vector (Fig. 1.1). When they are taken in a blood meal by the tsetse fly, they establish in the midgut and then migrate to the salivary glands to be transmitted to a new mammalian host where they live freely in the bloodstream. To complete the cycle, *T. brucei* differentiates into 5 distinct forms described below.

### **Procyclic form (PF)**

Procyclic forms are proliferative trypanosomes found in the midgut of the tsetse. VSG expression is blocked and the parasite's coat is less dense, made of GPI anchored procyclins. The mitochondrion is elongated extending through both ends of the cell and is fully active (Matthews 2005).

### **Epimastigote**

As the procyclics migrate and attach to the salivary gland wall by their flagellar membrane, they become epimastigotes and continue multiplying (Matthews 2005).

## **Metacyclic**

The proliferative epimastigotes undergo division arrest, start expressing VSG and are released into the salivary gland lumen, where they are identified as metacyclics (Matthews 2005).

## **Bloodstream form (BF)**

Through the tsetse fly bite, metacyclics are transmitted to a mammalian host, differentiate into BF, migrate to blood vessels, and live freely in the bloodstream. Initially, BF trypanosomes are long and slender (LS), rapidly dividing, with VSG expression positive and repressed mitochondrial activity. The LS mitochondrion is smaller in size, devoid of cristae and energy generation depends on glycosomes. As the parasite load increases in the blood, proliferative LS forms differentiate to nonproliferative shorter cells called short stumpy (SS) and arrest in G1 phase, maintaining positive VSG expression (Matthews 2005). SS forms depend on glycosome glycolysis but have partially functional mitochondrial system to allow survival in the fly midgut (Michels et al. 2006). As the number of SS forms increase in the bloodstream, the high density of SS parasites has been shown to inhibit LS proliferation, a mechanism to regulate the parasite population and prolong host survival (Seed and Wenck 2003). Arrest in G1 phase is necessary to guarantee successful completion of the cell cycle in the tsetse fly after they are taken up in a blood meal (Matthews 2005).

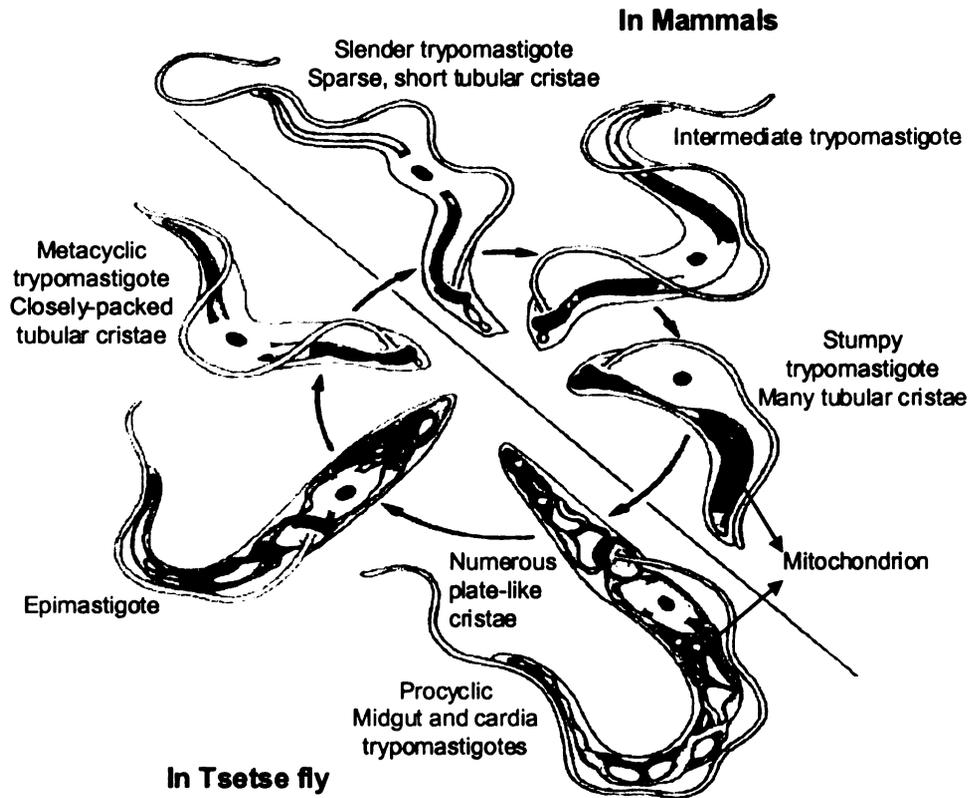


Figure 1.1 *T. brucei* life cycle through the mammalian (37°C) and insect (27°C) hosts. In the tsetse fly, the mitochondrion is large and reticulated, and ATP is produced through electron transport and oxidative phosphorylation. In contrast, during the bloodstream stage of the life cycle, the mitochondrion has a compact morphology and is less active; trypanosomes derive most energy from glycolysis. Reference: Vickerman, K. 1971. In Fallis, A.M., ed. Ecology and physiology of parasites. University of Toronto Press, Toronto.

In vitro studies utilize PF and BF cells because they can be axenically cultured (doubling time of 6-10 h) (Cunningham and Honigberg 1977; Hirumi and Hirumi 1989) and are amenable to gene function investigations through several genetic tools. The release of the *T. brucei* genome sequence (Berriman et al. 2005) is providing faster discoveries on gene function and understanding of trypanosome biology at the molecular, biochemical and cellular levels. Reverse and forward genetic techniques like transcription of stable transfection vectors (driven by endogenous RNA polymerase when integrated into the chromosome) and transient transfection vectors (endogenous RNA polymerase I or heterologous phage T7 promoters), drug resistance markers, homologous

gene replacement, conditional knockouts of essential genes by control of tetracycline inducible system, and RNA interference (RNAi) are available for *T. brucei*. Genetic and sexual exchanges can occur in the insect phase of *T. brucei* (Gibson et al. 2008) but limited availability of tsetse fly colonies and diploidy make it difficult to utilize some forward genetic techniques like negative drug selectable markers (Beverley 2003).

### **Cell and metabolism**

The trypanosome cell (10-30  $\mu\text{m}$ ) is elongated and maintained by a microtubule cytoskeleton. The cell contains a single copy of some organelles, such as nucleus, mitochondrion, kinetoplast, flagellum, and flagellar pocket.

### **Flagellar pocket**

The flagellar pocket is an invagination of the membrane at the posterior end of the cell that takes up 5% of the cell surface area. This pocket is the endo- and exo-cytosis site and where the flagellum exits the cell. In BF it has fast cell membrane uptake allowing complete VSG exchange in 12 min (Matthews 2005).

### **Flagellum**

The *T. brucei* flagellum is similar to a conventional eukaryotic axoneme associated with an extra-axonemal paraflagellar rod (PFR) and attached to the basal body via filaments associated with microtubules (flagellum attachment zone). The flagellum originates in the basal body, at the flagellar pocket, and has the function of promoting cell motility. PFR is a network of filaments attached along the length of the flagellum and assures proper motility. The basal body is linked to the kinetoplast through the mitochondrial membrane by a series of filaments that cross the cell and the mitochondrial membranes (Maga and LeBowitz 1999).

## **Glycosome**

Approximately 65 single-membrane surrounded glycosomes (0.2 to 0.3  $\mu\text{m}$  diameter) can be found in *T. brucei*. They do not contain DNA and are related to peroxisomes, which are organelles containing hydrogen peroxide-generating oxidases and catalase in yeast. In *T. brucei*, glycosomes lack peroxidases and catalases and some of their enzymes have low or no activity like enzymes for ether-lipid biosynthesis,  $\beta$ -oxidation of fatty acids, pentose-phosphate pathway, purine salvage, and biosynthetic pathways for pyrimidines (Hellemond et al. 2005). Unlike yeast peroxisomes, glycosomes are essential for *T. brucei* survival because they compartmentalize most of the cell's glycolytic enzymes that would be toxic if not in the glycosome. Together with the cytosol and mitochondrion, glycosomes are responsible for energy metabolism. The enzymatic content in the glycosome varies depending on the parasite's stage in the life cycle, allowing fast adaptation to different environmental conditions. For instance, in BF, both the aerobic and the anaerobic glycolysis take place in glycosomes (Fig. 1.2) utilizing glucose abundantly present in the blood and body fluids like the cerebrospinal fluid (Michels et al. 2006). From glucose, 3-phosphoglycerate (3-PGA) is produced in the glycosome, then shuttled to the cytoplasm to be converted to pyruvate (which is excreted) with production of ATP. To maintain glycosomal redox balance under aerobic conditions, NADH generated by glycerol-3-phosphate (Gly-3-P) dehydrogenase is reoxidized by a mitochondrial glycerol-3-phosphate shuttle and returned to the glycosome as dihydroxyacetone phosphate (DHAP) to cyclically produce more Gly-3-P and  $\text{NAD}^+$  (Parsons 2004). Under anaerobic conditions the redox balance is maintained by degrading glucose into equal amounts of glycerol and pyruvate; however, glycerol

production is thermodynamically unfavorable making anaerobic glycolysis insufficient for growth (Hellemond et al. 2005).

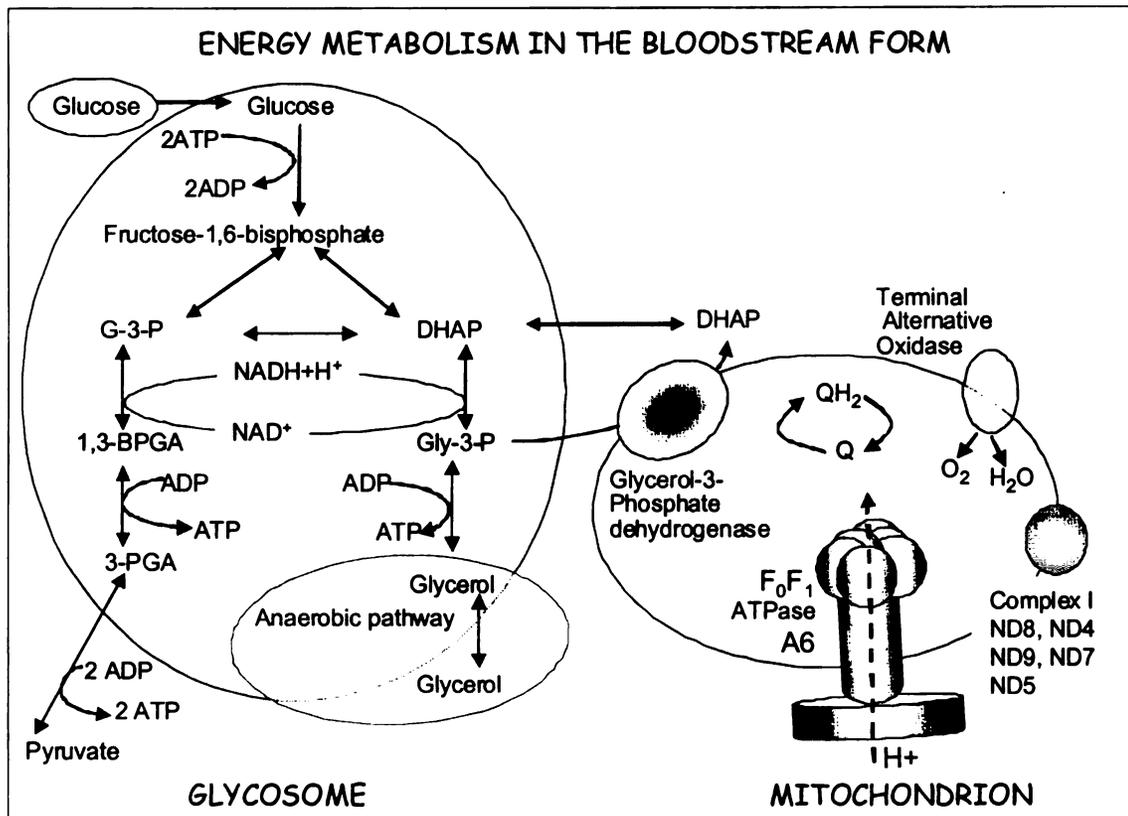


Figure 1.2 Schematics of energy metabolism in the bloodstream form *T. brucei*. In this stage, glucose is used as the main energy source. The first seven steps of glycolysis and much of the pentose phosphate pathway occurs in the glycosome. The last three steps of glycolysis are cytoplasmic after 3-phosphoglycerate is shuttled to the cytoplasm, generating two net ATP. The mitochondrion in the bloodstream form is largely repressed. A putative glycerol 3-phosphate/DHAP shunt to the mitochondrion is thought to help maintain NAD<sup>+</sup>/NADH balance via glycerol 3-phosphate dehydrogenase and the terminal alternative oxidase. Transcription and editing of some of the mRNAs from Complex I of the respiratory chain (ND4, 5, 7-9) are upregulated in this stage. A F<sub>0</sub>F<sub>1</sub>-ATPase maintains an electron membrane potential important for survival. A6 (ATPase subunit 6) is edited in this stage. BPGA: 1,3-bisphosphoglycerate, DHAP: dihydroxyacetone phosphate, G-3-P: glyceraldehyde 3-phosphate, ND: NADH dehydrogenase, 3PGA: 3-phosphoglycerate, QH<sub>2</sub>/Q: ubiquinol/ubiquinone. Figure by Melissa K. Mingler.

The tsetse fly respiratory system provides oxygen to all its tissues; therefore, aerobic metabolism (Fig. 1.3) predominates in insect stages (Michels et al. 2006). PF trypanosomes are very flexible in their metabolism, for instance, glucose is scarce in tsetse because the blood taken by the fly is rapidly degraded to amino acids

(predominantly proline and threonine). Therefore, the main energy source for the PF trypanosome are the aminoacids, which are metabolized in the mitochondrion. On other hand, if there is any glucose available to the PF trypanosome, the glucose will be consumed and catabolized to succinate in the glycosome (a pathway that includes few steps in the cytoplasm) (Fig. 1.3). Succinate is additionally produced in the mitochondrion, utilizing the malate produced in the glycosome. The production of succinate in the the glycosome and in the mitochondrion of the PF trypanosome is sufficient to maintain the glycosomal and the glycolytic redox balance. Therefore, the Gly3P/DHAP shuttle (necessary for the BF trypanosome, described above and below, under “mitochondrial metabolism”) is unimportant for the PF cell, although existent if necessary. Having a functional shuttling system provides metabolic flexibility to the PF, which may be important in case of stress in the growth conditions. Another exemple of PF metabolic flexibility was found when it was cultivated in a glucose-rich media. In this environment, glucose exerts a negative pressure on proline metabolism and glycosomes catalyze glucose somewhat similarly to BF (with some steps in the cytosol). In this case, pyruvate is generated in the cytosol by conversion of phosphoenolpyruvate by cytosolic pyruvate kinase. The cytosolic pyruvate kinase has been shown to be essential for PF survival, especially because this reaction produces ATP and the pyruvate is further converted to acetyl CoA, which feeds the tricarboxylic acid cycle, generating more ATP (Coustou et al. 2003). In absence of glucose, glycolysis is dispensable in PF and expression of glycolytic enzymes is reduced (van Weelden et al. 2005).

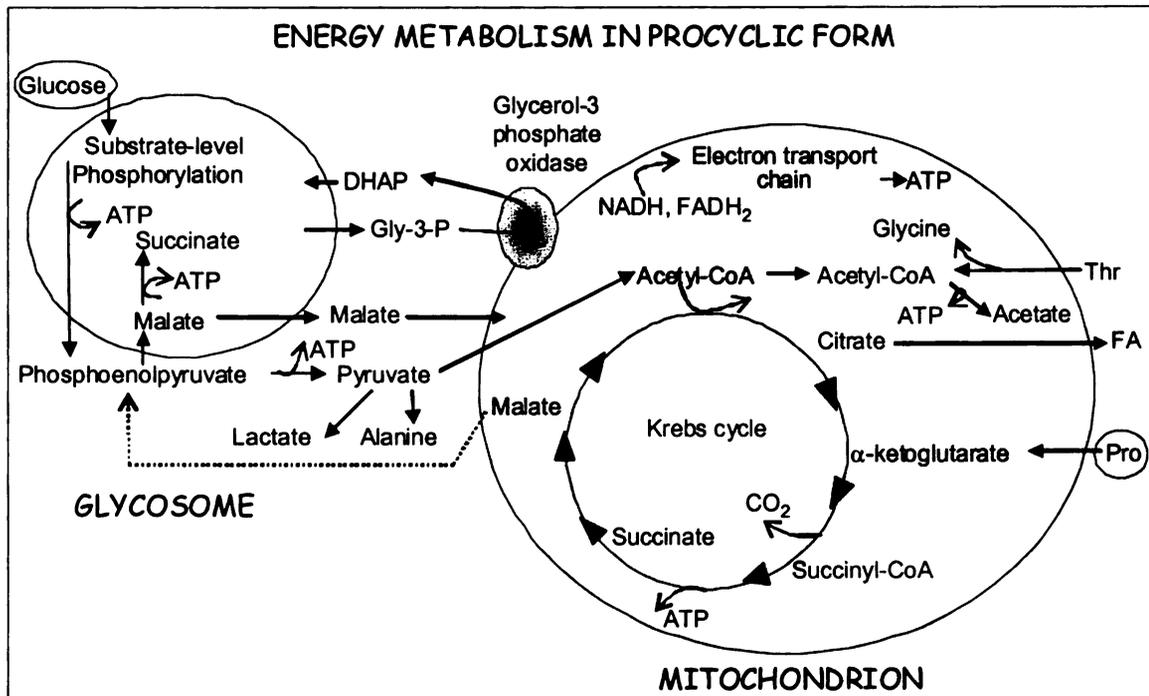


Figure 1.3 Schematics of energy metabolism in the insect stage, or procyclic form *T. brucei*. With scarce glucose supply, the glycosome functions like anaerobic organisms with succinic fermentation to maintain energy production. The Krebs cycle (divided in 3 highlighted parts) and respiration in the mitochondrion utilize amino acids [mainly proline (Pro) and threonine (Thr)] and any glucose available as energy sources. DHAP: dihydroxyacetone phosphate, Gly-3P: glycerol 3-phosphate, FA: fatty acids. Figure by Melissa K. Mingler.

### Mitochondrion (mt)

Like any eukaryotic mitochondria, the *T. brucei* mitochondrion (there is only one mt per cell) has a double lipid bilayer membrane, contains DNA, and works in oxidative phosphorylation. It differs from the mitochondrion of other eukaryotes in morphology, genome, and biogenesis. The trypanosome mitochondrial DNA is organized into a disk-like structure called the kinetoplast (k) or kDNA (Matthews 2005). The kDNA encodes 18 proteins, 2 very small ribosomal subunits and 2 guide RNAs. All the other mt proteins and tRNAs are nuclear encoded and must be imported from the cytosol (Lukes et al. 2005). Mitochondrial gene expression includes polycistronic transcription, cleavage to release monocistrons, mRNA editing if necessary, polyadenylation of mRNAs,

polyuridylation of ribosomal (r) and guide (g) RNAs, and translation (Lukes et al. 2005). The mt activities vary according to the trypanosome life cycle. BF has a mostly repressed mt while oxidative and substrate level phosphorylation only occur in PF.

### **Mitochondrial metabolism**

The long slender BF in the mammalian host generates energy (ATP) through glycolysis in the glycosome and cytoplasm. Its mitochondrion has a simple morphology, lacks cristae, lacks most Krebs cycle enzymes, the cytochrome-containing respiratory complexes, and a classical respiratory chain, but contains a plant like alternative oxidase. This alternative oxidase is important to maintain the glycosomal and glycolytic redox balance through the Gly-3-P, dihydroxy-acetone phosphate shunt (Besteiro et al. 2005). Reducing equivalents (NADH) produced in the glycosomes are transferred to the mt via a classical mammalian type shuttle and the electrons donated to the ubiquinone/ubiquinol pool. The reduced ubiquinol is the electron donor for the plant-like alternative oxidase. This does not involve  $H^+$  translocation and therefore does not drive ATP production (Fig. 1.2) (Hellemond et al. 2005). In BF, the  $F_0F_1$ -ATP synthase complex (respiratory complex V) is functional but at a lower expression level than in PF. The subunit 6 of the  $F_0F_1$ -ATP synthase is encoded in the kDNA, edited and expressed in both LS and PF stages. This complex is located in the inner mt membrane ( $F_0$  contains subunits 6 and 4, one of each, and several subunits 9 in a ring, all located in the membrane;  $F_1$  comprises a ring of six alternating  $\alpha$  and  $\beta$  subunits with the ATP synthase catalytic sites extending into the matrix). In PFs and in most eukaryotic organisms, protons generated by the respiratory proton pumps flow through a channel created by  $F_0F_1$  subunits, see figure 1.3, causing rotation and conformational changes of some subunits, resulting in ATP

synthesis (Schnauffer et al. 2005). The BF utilizes a reverse pathway where  $F_0F_1$ -ATP synthase hydrolyzes ATP to pump protons and generate mitochondrial membrane potential ( $\Delta\Psi_m$ ). This  $\Delta\Psi_m$  is absolutely required for mitochondrial import of nuclear proteins and other processes; thus, essential for mitochondrial activities and biogenesis. Diskinetoplastids (trypanosomes that lost their mitochondrial genome, see more under kinetoplast below) also need this reversal mechanism to survive and the existence of a mutation in  $F_1$  helps *T. evansi* evade the need for gene products encoded in the mitochondrion, like ATP synthase subunit 6 (Schnauffer et al. 2005).

Procyclic *T. brucei* contains a completely developed mitochondrion and a more complex energy metabolism than BFs. The PF energy metabolism depends mainly on degradation of pyruvate to acetate and of amino acids to succinate and/or acetate, as shown in figure 1.3 (Besteiro et al. 2005). Although PFs present and express all eight enzymes for the Krebs cycle, the parasite does not use it as a cycle. Procyclics have the flexibility of using a partitioned cycle, likely due to low expression of certain enzymes, or the cycle is supposedly divided in three to maintain the mt redox status (Fig. 1.3). One third of the cycle is catabolic (from  $\alpha$ -ketoglutarate to succinate) and is utilized for proline degradation. Another third of the cycle is proposed to be anabolic, where citrate is used for the biosynthesis of fatty acids. The last third of the Krebs cycle forms malate from succinate and the malate is utilized for the synthesis of phosphoenolpyruvate in the cytosol. This part of the cycle is also anabolic since succinate (from proline) can be used for production of acetate or for gluconeogenesis in the absence of glucose or glycerol (van Weelden et al. 2005). The above pathways produce ATP and also large amounts of  $FADH_2$  and NADH, which are reoxidized by the respiratory chain (Fig. 1.4). PF contains

homologues for classical mammalian respiratory chain complexes I-IV (used to reoxidize NADH by oxidative phosphorylation and generate ATP) and a plant-like alternative oxidase (Hannaert et al. 2003; Hellemond et al. 2005). Oxidative phosphorylation is essential in PFs; however, if they are grown in glucose-rich medium, enough ATP can be obtained from substrate-level phosphorylation (Coustou et al. 2003; Besteiro et al. 2005). In the respiratory chain, the electrons transferred to the ubiquinone pool come from complex I (NADH:ubiquinone oxidoreductase), complex II (succinate dehydrogenase) and glycerol-3-phosphate dehydrogenase. From the ubiquinol, the electrons are transferred to the alternative oxidase, complex III (cytochrome c reductase) and complex IV (cytochrome c oxidase). Only the last two complexes translocate protons resulting in a proton motive force used for ATP production by the ATP synthase (Hellemond et al. 2005).

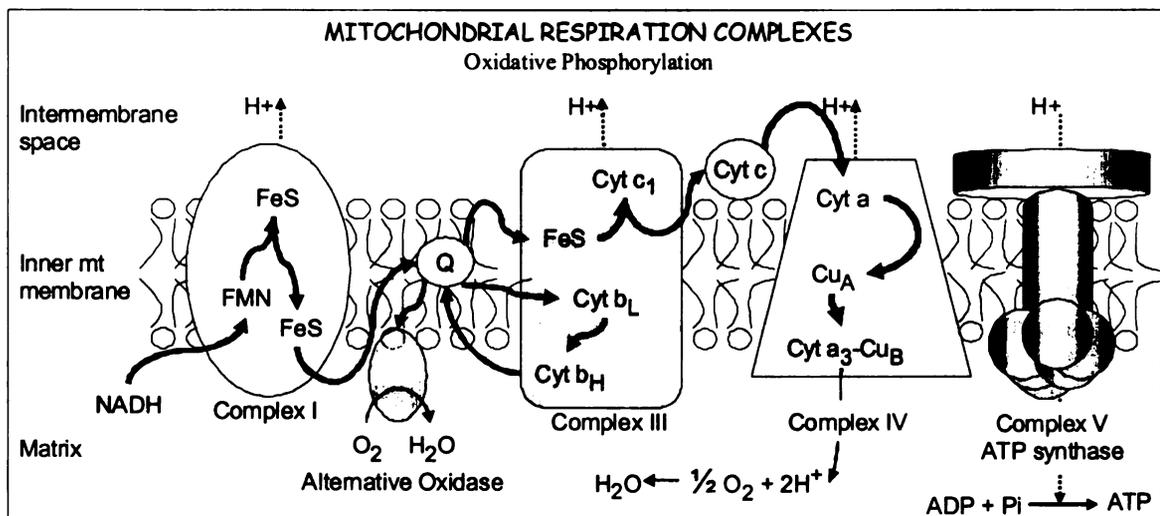


Figure 1.4 Schematics of the mitochondrial respiratory complexes in *T. brucei*. Complexes I, III, IV, and V are present as well as the alternative oxidase. Complex II is predicted to exist because a succinate dehydrogenase activity and succinate dependent respiration are present; however, the actual proteins for the complex have not been confirmed. Cyt = cytochrome, Cu = copper center, H<sup>+</sup> = proton, FeS = iron/sulfur cluster, FMN = flavin mononucleotide, NADH = Nicotinamide adenine dinucleotide in its reduced form, Q = ubiquinone. Figure by Melissa K. Mingler.

## **Kinetoplast (kDNA)**

The kinetoplast comprehends two classes of circular molecules of different sizes and content, the maxi- and minicircles, which are interlocked among themselves forming a  $10^7$  kDa dense network (Schneider 2001). The kDNA is fragile and can be easily altered or lost by intercalating drugs in the laboratory or in nature, creating induced and natural diskinetoplastic (Dk) trypanosomes. Although the mitochondrion is always present, most of the kDNA is usually missing in Dk parasites, and the amount of missing kDNA varies depending on the species and strain. Loss of all kDNA is called akinetoplastidy (Ak). Dk or Ak trypanosomes (e.g. *T. equiperdum* and *T. evansi*) may have originated from *T. brucei* (Lai et al. 2008). While they have lost the capability to differentiate into the insect stage, they are viable as BFs (maintain  $\Delta\Psi_m$  with  $\alpha$  and  $\beta$  subunits of ATP synthase) and are transmitted mechanically by biting flies or venereally (Ou et al. 1991; Schnauffer et al. 2002; Domingo et al. 2003). Despite the fact that Dk trypanosomes can be viable as BFs, altering kDNA replication or expression is lethal for still unknown reasons (Schnauffer et al. 2002).

## **Maxicircles**

A maxicircle is a circular DNA that resembles the mt DNA of other organisms. In trypanosomes, there are 20-50 identical maxicircle copies of approximately 22 kb each copy (Fig. 1.5). They encode typical mitochondrial gene products such as six subunits of complex I, apocytochrome b of CoQ-cytochrome c reductase (complex III), three subunits of cytochrome oxidase (complex IV), subunit 6 of ATP synthase (complex V), a ribosomal protein (RPS12), two ribosomal RNA genes (9S and 12S), and two gRNAs (Figs. 1.4, 1.5). Six open reading frames remain unidentified, maxicircle unidentified reading frame (MURF) 1, 2, and 5, and CR3, 4 and 5 (GC-rich sequences) (Simpson et

al. 1998). The striking difference is on some of the protein-coding genes that are encrypted (note the overlapped genes in figure 1.5) and need to undergo posttranscriptional modification via RNA editing to be translatable (Simpson 1987).

### **Minicircles**

Minicircles are also circular DNA, like the maxicircles, but smaller in size, 0.9 to 2.5 kb (1 kb in *T. brucei*), heterogeneous in sequence, and exist in higher number (5,000-10,000). In *T. brucei*, two or three gRNAs are encoded in a minicircle and are flanked by 18 bp inverted repeats proposed to work in transcription because transcription begins 32 bp from the upstream repeat (Pollard et al. 1990). Guide RNAs are mostly primary transcripts, including the maxicircle encoded gRNAs, but transcription can be polycistronic (Hajduk and Sabatini 1998; Clement et al. 2004).

Figure 1.5. Schematics of *T. brucei* maxicircle, minicircle, gRNA structure, editing progression and editing steps. **A.** Maxicircle: composed by a conservative region (17 kb) indicated by gene acronyms and a divergent region of repeated sequences of variable structure. **B.** Minicircle and relative position of gRNA genes. **C.** gRNA domains: anchor, guiding region and U-tail. **D.** Editing progression 3' to 5' within pre-edited mRNA generates edited region and can create binding site for next gRNA. **E.** Insertion and deletion editing steps. The mRNA is cleaved by an endonuclease at the editing site (ES), U(s) are inserted by a terminal uridylyl transferase (TUTase) or deleted by a exoribonuclease (exoUase), and the mRNA fragments are joined by a ligase. Acronyms: 12S, lsrRNA; 9S, ssrRNA; ND1,4,5,7-9, NADH-dehydrogenase subunits; CO1-3, cytochrome oxidase subunits; Cytb, cytochrome b of the bc1-complex; A6, subunit 6 of the adenosine triphosphatase; MURF, maxicircle unidentified reading frame; RPS12, ribosomal protein S12; G, gRNA; CR, C-rich regions. Figure adapted from Simpson et al (1998) and Stuart et al (2005).

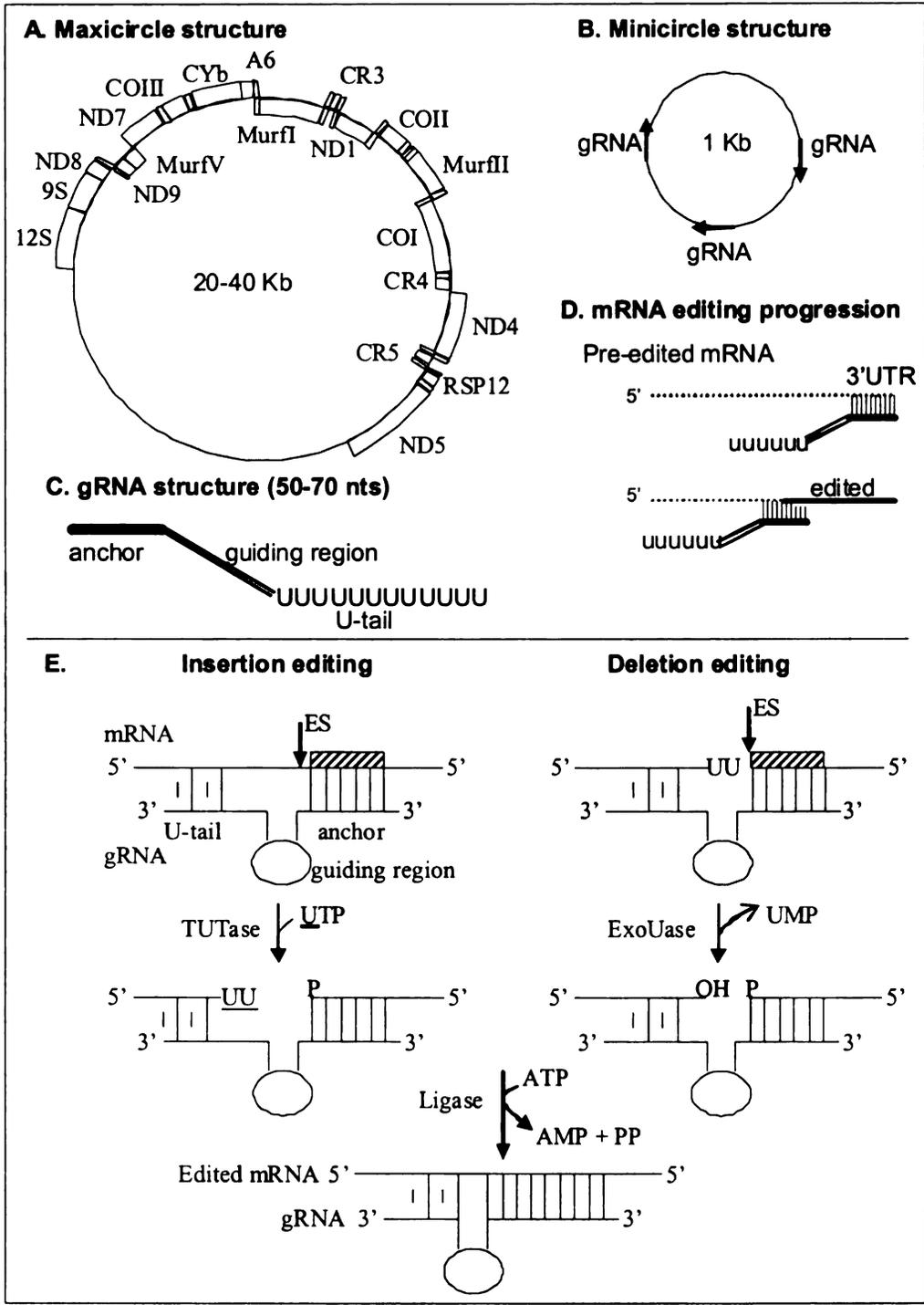


Figure 1.5 Schematics of *T. brucei* maxicircle, minicircle, gRNA structure, editing progression, and editing steps.

## ***RNA editing***

### **Control of mitochondrial biogenesis**

Initial genomic DNA sequencing of mitochondrial DNA from the trypanosomatids indicated that many typical genes seemed to be lacking or defective. However, sequencing of cDNA indicated that translatable open reading frames (ORF) for 12 of the 20 genes are created posttranscriptionally through RNA editing (Benne et al. 1983; Benne et al. 1986; Benne 1994). Editing is precise and involves insertion of hundreds of uridylates (U)s and less frequently deletion of Us, often doubling the transcript size (Feagin et al. 1988; Koslowsky et al. 1990). The modifications repair frameshifts, generate start codons, and create entire ORFs (Benne et al. 1986; Benne 1994; Seiwert and Stuart 1994; Simpson and Thiemann 1995; Lukes et al. 2005). Editing can also create alternative distinct ORFs (Ochsenreiter and Hajduk 2006; Ochsenreiter et al. 2008). This type of processing is unique to trypanosomatids and essential as it allows proper mitochondrial function (Schnauffer et al. 2002; Schnauffer et al. 2005). Editing extent and control vary depending on the mRNA (table 1.2) and organism (Hajduk and Sabatini 1998; Simpson et al. 1998); therefore, many of the proteins for energy metabolism that are translated in the mitochondrion are developmentally regulated by RNA editing. In BF *T. brucei*, many of the mitochondrial respiratory complex proteins including cytochrome c oxidase and cytochrome c reductase complexes are down regulated by editing. Apocytochrome b (CYb) is edited in PF only, at a small segment near its 5' end (Feagin et al. 1987). NADH dehydrogenase (ND) subunits 8 and 9 are low in PF trypanosomes but increase in the BF (Koslowsky et al. 1990; Souza et al. 1992). ND7 is differentially and extensively edited in two separate domains; in procyclics only the 5' domain is edited whereas in bloodstream the transcript is fully edited (5' and 3'

domains) yielding two forms of the mRNA (Koslowsky et al. 1990; Souza et al. 1992; Souza et al. 1993; Simpson et al. 1998; Koslowsky 2004). Developmental regulation by RNA editing controls the expression of mitochondrial proteins, and consequently, regulates mitochondrial biogenesis throughout the trypanosome life cycle. Very little is known of how editing is regulated. A few mechanisms for mRNA editing regulation have been investigated but its true mechanism remains to be elucidated. It was thought that gRNA abundance could control editing but gRNAs for developmentally regulated messages are present in both PF and BF (Koslowsky et al. 1992; Riley et al. 1994; Riley et al. 1995). Regulation of transcript abundance has been correlated with cleavage of polycistronic pre-mRNAs and polyadenylation. These post-transcriptional processing events seem to cross-talk with editing; for instance, addition of long poly(A/U) tails only happens after completion of editing, editing has been observed in overlapping regions between two still uncleaved messages, and 3' cleavage of COII mRNA only occurs after completion of editing (Koslowsky and Yahampath 1997; Militello and Read 1999; Etheridge et al. 2008). Although regulation is still obscure, these posttranscriptional processing events are independent but able to influence one another (Etheridge et al. 2008).

Table 1.2 Mitochondrial editing modification for each of the *T. brucei* mRNAs.

Gene	No. of uridines		Edited size (nt)	Life stage edited
	Added	Deleted		
CYb	34	0	1,151	Pro <sup>a</sup>
A6	448	28	821	Pro/Bs <sup>b</sup>
COI	0	0		Unedited <sup>c</sup>
COII	4	0	663	Pro
COIII	547	41	969	Pro/Bs
ND1	0	0		Unedited
ND3	210	13	452	Unknown
ND4	0	0		Unedited
ND5	0	0		Unedited
ND7	553	89	1,238	5'Pro/Bs, 3'Bs <sup>d</sup>
ND8	259	46	574	Bs
ND9	345	20	649	Bs
S12	132	28	325	Bs
MURF1	0	0		Unedited
MURF2	26	4	1,111	Pro/Bs
CR3	148	13	299	Bs
CR4	325	40	567	Bs

<sup>a</sup> Pro, transcript is edited only in the procyclic developmental stage.

<sup>b</sup> Pro/Bs, transcript is edited in both bloodstream and procyclic stages.

<sup>c</sup> Editing of these transcripts has not been reported.

<sup>d</sup> The ND7 transcript is differentially edited in the procyclic and bloodstream developmental stages.

Reference: Hajduk and Sabatini (1998).

### Mitochondrial-encoded Adenosine triphosphatase subunit 6 (A6)

Most of the work presented here focused on the study of *T. brucei* A6 mRNA (earlier MURF 4) and its 3'-most initiating gRNA, gA6-14. The A6 protein is a component of the F<sub>0</sub>F<sub>1</sub>-ATP synthase mitochondrial complex (respiratory complex V) and is essential for survival of all life stages of the parasite (Schnauffer et al. 2005). Within the kDNA, the A6 gene is located between Cyb and MURF1; it is GC rich with G versus C strand bias before and after editing (Figs. 1.5, 1.6). This G bias is a characteristic of extensively edited transcripts, such as COIII (Feagin et al. 1988; Bhat et al. 1990). The entire A6 transcript is constitutively and extensively edited, 448 Us are inserted at 173 sites and 28

encoded Us are removed from 12 sites, resulting in a transcript with 821 nucleotides, excluding the poly(A) tail. This extensive editing creates the A6 ORF, also start and stop codons. For unknown reason, the A6 transcript curiously has a large (143 nt) 3' untranslated region (UTR) that is edited. Differently from *T. brucei*, the A6 gene from *L. tarentolae* is larger and only needs extensive editing at the 5' end (Fig. 1.6) (Bhat et al. 1990). gA6-14 base pairs within A6's 3' end and directs editing of 34 editing sites, with a total of 10 deletions and 81 insertions (Bhat et al. 1990; Seiwert and Stuart 1994; Leung and Koslowsky 2001b).

Figure 1.6 DNA and fully edited mRNA sequences of the ATPase subunit 6 from *T. brucei* and *L. tarentolae*.

Lower case uridines (u) were inserted by RNA editing and the deleted Us are represented by asterisks. Underlined are the start and stop codons created by editing. Figure from Bhat et al (1990).





## **gRNAs**

To guide the process of editing there are more than 1200 mitochondrially encoded gRNAs in *T. brucei* (Stuart et al. 2005). They are small transcripts (50-70 nt) that contain a distinctive primary structure, see figure 1.5 (Blum et al. 1990; Byrne et al. 1996; Kable et al. 1996). The 5' most part, or anchor, defines the specificity of the gRNA, it is 4-18 nt in length and complementary to the unedited or partially edited mRNA. The first mismatched nucleotide pair, 3' of the anchor duplex formed between gRNA and mRNA, indicates the first ES and the beginning of the guiding region. The central guiding region provides the information necessary for U-insertion and deletion on the mRNA (Adler and Hajduk 1997; Simpson 1997). The 3' end of the gRNA is the oligo U-tail (5-24 Us) added by a terminal uridylyl transferase (TUTase, KRET1) (Blum and Simpson 1990; Aphasizhev et al. 2002). The U-tail role is still unclear. It is known to interact with the purine-rich pre-edited region, likely stabilizing the mRNA/gRNA complex; it may also tether the 5' mRNA fragment within the editing complex (Blum and Simpson 1990; Seiwert et al. 1996; Leung and Koslowsky 1999; 2001b; Koslowsky 2004) and may help in earlier stages of the gRNA/mRNA hybridization pathway (Yu 2006).

## **Editing machinery (editosome)**

The editing machinery is composed of many proteins and enzymes with several nomenclatures (Stuart et al. 2005; Panigrahi et al. 2006; Law et al. 2007; Carnes et al. 2008). It has been purified from *T. brucei* and *L. tarentolae* by several labs (Pollard et al. 1992; Rusche et al. 1997; Madison-Antenucci and Hajduk 2001; Aphasizhev et al. 2003; Panigrahi et al. 2007). The latest results confirm the presence of at least three distinct editosomes (characterized mainly through tandem affinity purification of editosome proteins) proposed to work in conjunction to edit the mitochondrial transcripts (Panigrahi

et al. 2006; Panigrahi et al. 2007; Carnes et al. 2008). The repression of editing genes by either conditional inactivation of expression or RNAi using tetracycline regulatable promoters revealed that almost all editosome genes are essential to survival of the trypanosomes (Schnauffer et al. 2001; Wang et al. 2003; Carnes et al. 2005; Trotter et al. 2005).

Each editosome has a common set of proteins and one or two specific proteins that make them functionally distinct. They all contain at least 20 proteins (table 1.3), sediment at ~20 Svedberg (20S) on glycerol gradients, and each has three editing activities (observed by in vitro RNA editing assays):

1) Endonuclease. TbMP90 (KREN1) cleaves the mRNA at U-deletion sites or TbMP61 (KREN2) cleaves at U-insertion sites (Carnes et al. 2005; Trotter et al. 2005). A third endonuclease, TbMP67 (KREPB2/KREN3), cleaves at insertion sites when the gRNA is found in cis (Panigrahi et al. 2006; Carnes et al. 2008).

2) TUTase or exoUase. The terminal uridylyl transferase (TUTase) TbMP57 (KRET2) inserts U(s) to the 3' end of the 5' cleavage fragment (Aphasizhev et al. 2003; Ernst et al. 2003). Several U-specific exoribonuclease (exoUase) responsible for deleting uridines have been identified. These include TbMP99 (KREPC2/KREX2), TbMP100 (KREPC1/KREX1) and TbMP42 (KREPA3; band-VI) (Schnauffer et al. 2003; Worthey et al. 2003; Brecht et al. 2005; Kang et al. 2005).

3) Ligase. Two RNA ligases have been identified; TbMP52 (KREL1; band-IV) is responsible for ligation at U-deletion sites, while generally TbMP48 (KREL2; band-V) ligates U-insertion sites (Rusche et al. 1997; Huang et al. 2001; McManus et al. 2001; Rusche et al. 2001; Schnauffer et al. 2001; Cruz-Reyes et al. 2002; Schnauffer et al. 2003).

Table 1.3 List of *T. brucei* editosome proteins isolated from U-insertion and U-deletion editosomes, their function and functional domains.

Protein	U ins./deletion	Function	Functional domains
KREPA1	both	Interaction/U-ins. Recognition	1 zinc finger, 1 zinc-like finger, OB fold
KREPA2	deletion	Interaction	2 zinc finger, OB fold
KREPA3	both	Interaction/cleavage factor	2 zinc finger, OB fold
KREPA4	both	Interaction	OB fold-like
KREPA5	both	Interaction	OB fold-like?
KREPA6	both	Interaction	OB fold-like
KREN1	deletion	endonuclease	U1-like zinc finger, RNaseIII, dsRBD
KREN3 (KREPB2)	insertion	endonuclease	U1-like zinc finger, RNaseIII, dsRBD
KREN2	insertion	endonuclease	U1-like zinc finger, RNaseIII, dsRBD
KREPB4	both	Interaction	U1-like zinc finger, RNaseIII-like, Pumilio
KREPB5	both	Interaction	U1-like zinc finger, RNaseIII-like, Pumilio
KREPB6	*	Interaction	U1-like zinc finger
KREPB7	insertion	Interaction	U1-like zinc finger
KREPB8	deletion	Interaction	U1-like zinc finger
KREX1	deletion	exoUase	5'3' exonuclease, endo, exo, phosphatase
KREX2	both	exoUase	5'3' exonuclease, endo, exo, phosphatase
KREL1	both	RNA ligase	Ligase, tau (microtubule assoc.), kinesin light chain
KREL2	insertion	RNA ligase	Ligase, tau (microtubule assoc.), kinesin light chain
KRET2	insertion	TUTase	Nt transferase, core, poly(A)polymerase associated domain
KREH1	transient	Helicase	Helicase with DEAD box
<b>KRET1 complex</b>			
KRET1	?	gRNA TUTase	zinc finger, poly(A)polymerase catalytic and associated domains
<b>Accessory proteins</b>			
<b>MRP complex</b>			
MRP1	?	RNA matchmaking	R-rich domain
MRP2	?	RNA matchmaking	R-rich domain
<b>Others</b>			
RBP16	?	RNA annealing	Cold shock, RGG RNA binding domain
REAP-1	?	Interaction	21 AA repeat
TbRGG1	?	Interaction	RGG RNA binding domain

\* KREPB6 was associated with KREN3 editosome only.

References: Stuart et al (2005), Carnes et al (2005), Panigrahi et al (2006), Carnes et al (2008).

The editosomes are mainly distinguished by the different endoribonucleases, KREN1-3 (Cruz-Reyes et al. 1998; Carnes et al. 2005; Trotter et al. 2005; Carnes and Stuart 2007). Each of the three editosomes contains only one of the three KREN. KRENs have RNase III, U1-like Zn<sup>+2</sup> finger (may facilitate interaction with RNA or other proteins), and dsRNA binding motifs. Editosomes containing KREN1 are specific

for cleaving insertion sites, and the ones containing KREN2 cleave only at deletion sites. Editosomes containing KREN3 cleave insertion sites with specificity for COII pre-mRNA (Carnes et al. 2008). The other RNase III-like-containing proteins are KREPB4 and 5, which are present in all three editosomes. While they are critical for editosome integrity, their specific function remains to be elucidated (Wang et al. 2003; Panigrahi et al. 2006). Repression of KREPB4 or 5 dramatically affects editosome function and integrity; KREPB4 knock-downs reduce insertion and deletion site endonuclease activities thus this protein could be important for this initial step of editing (Babbarwal et al. 2007). KREPB4 and 5 may also form heterodimers with KREN1-3 (Carnes et al. 2008) since RNase III endoribonucleases usually exist as homodimers (Nicholson 1999).

The ligases KREL1 and 2 are more likely to exist in all editosomes. KREL1, but not KREL2, is essential and can compensate for the absence of KREL2, but not vice-versa. KREL1 was found to be active in both insertion and deletion systems while KREL2 only ligates in insertion assays (Cruz-Reyes et al. 1998; Cruz-Reyes et al. 2002; Palazzo et al. 2003; Schnauffer et al. 2003).

Six KREPA proteins have been identified with an OB-fold motif that facilitates binding to nucleic acids. KREPA1-3 additionally have N-terminal zinc fingers. KREPA1 and 2 were found to closely interact with KREL2 and 1 respectively, possibly providing the OB-fold in trans to the ligases, given that KRELs do not have a C-terminal OB-fold like other DNA ligases (Schnauffer et al. 2003). KREPA2, but not KREPA3, is a potential protein that crosslinks to ES1 for U-deletion in A6 pre-hybridized with a gRNA (D33, which lacks most of the guiding nucleotides and has a modified U-tail) within the editosome (Sacharidou et al. 2006). KREPA3 (TbMP42) is an interesting RNA-binding

protein; its recombinant version exhibited endonuclease activity even though the protein lacks canonical nuclease domain. In addition, it showed 3'-U-exo activity that could be used in regular editing or to remove the excess U's added by TUTase (Igo et al. 2002; Brecht et al. 2005). RNAi knockdown of KREPA3 does not disassemble the editosome but reduces in vivo and in vitro editing (Brecht et al. 2005). In contrast, recent data shows that KREPA3 improves editosome stability and does not have catalytic activity, yet it is required for the endonuclease cleavages, likely by facilitating substrate recognition (Law et al. 2008). KREPA4 and 6 are RNA binding proteins, with preference for stretches of U, and have critical roles in the structural integrity of ~20S editosomes as they establish protein-protein interactions (Salavati et al. 2006; Tarun et al. 2008).

The other structural proteins, KREPBs, promote protein-protein and protein-RNA interactions important for proper editosome assembly and stability (Stuart et al. 2005). KREPB6 was isolated exclusively from KREN3 ~20S editosome, KREPB7 from KREN2, and KREPB8 from KREN1 editosomes. These KREPBs may be associated with KRENs to add specificity in recognizing the distinct editing sites (Carnes et al. 2008).

Regarding KREXs, it is still unclear but it is likely that KREX1 catalyzes the removal of U's from deletion sites; while KREX2 has a more alternative role possibly trimming excess U's (e.g., added by KRET2) as it is present in both KREN1 and KREN2 complexes (Panigrahi et al. 2006). Finally, a helicase KREH1 was found in 20S editosomes isolated by column chromatography and immunoprecipitation but it is not a stable component of the complex since it is absent from 20S preparations purified by tandem-affinity purification (TAP) procedures (Missel et al. 1997; Panigrahi et al. 2003).

KREH1 is not essential for *T. brucei* and its loss of function may be compensated by a second putative helicase not yet characterized (Panigrahi et al. 2003).

Extensive work has been conducted to decipher the editosome components, their individual function, and overall structural organization. The relative stoichiometry of the proteins, structural organization, and how the machinery works remains to be determined. The protein stoichiometry may vary between cell lines, as there was discrepancy on the number of proteins isolated from the editosome between different labs.

### **Accessory proteins**

While the editosome is responsible for catalyzing the editing process, other mitochondrial proteins and complexes not associated with the editosome are also important in the editing process (table 1.3). The KRET1 complex adds the oligo U-tail to gRNAs (Aphasizhev et al. 2003). The MRP complex (MRP1 or gBP21 and MRP2 or gBP25) is involved in mRNA/gRNA annealing and may play additional roles in RNA metabolism (Lambert et al. 1999; Muller et al. 2001; Muller and Goring 2002; Aphasizhev et al. 2003; Vondruskova et al. 2005; Schumacher et al. 2006). Other RNA-binding proteins including RBP16 were shown to enhance in vitro insertion editing of CYb and promote overall mRNA stability (Pelletier and Read 2003; Miller et al. 2006). For both RBP16 and the MRP complex, the effect on RNA editing and stability appears to be transcript specific, with opposite results found for different RNAs. REAP1 and TbRGG1 have uncertain functions. REAP1 binds to polypurine tracks within pre-mRNAs and may have a role in editing (possibly bringing the mRNA to the editing complex) and stability. Loss of REAP-1 in single-knockout cell lines leads to a global increase in unedited, edited, and never-edited RNAs, confirming its role in RNA metabolism (Madison-Antenucci and

Hajduk 2001; Hans et al. 2007). TbRGG1 is an RGG (Arg-Gly-Gly) type protein shown to bind oligo-U sequences (Vanhamme et al. 1998) and appears to function in stabilizing edited RNAs or increasing editing efficiency (Hashimi et al. 2008). While the presence of these proteins is unnecessary for in vitro editing assays, they may be required in vivo for efficient editing of certain complexes or for mRNA/gRNAs with weak binding affinities and/or association rate constants.

### **RNA editing mechanism**

The development of RNA editing assays in vitro allowed elucidation of basic molecular mechanisms. The “full-round” editing assay (Seiwert and Stuart 1994) takes mitochondrial extracts or biochemically purified editosomes combined with a radiolabeled substrate mRNA and its cognate gRNA, and can access all editing reactions, the endonucleolytic cleavage, U insertion or deletion, and ligation.

A drawback of the full-round in vitro system is the formation of dead-end products or chimeras by a covalent linkage of the gRNA 3' end with the 5' end of the 3' mRNA fragment (Blum et al. 1991). These were found to be aberrant products that are also detected in vivo but in low levels (Riley et al. 1995). Another limitation is the inefficiency of in vitro systems (3-5% of the molecules are edited), due to inefficient endonucleolytic cleavage, and only a single round of editing was detected (Stuart et al. 2004). The efficiency of in vitro editing assays was increased after development of a “precleaved” system, where two short fragments of the mRNA (representing the 3' and 5' fragments after the endonucleolytic cleavage step) were annealed to an gRNA bridge to bypass the limiting endonucleolytic cleavage step (Igo et al. 2000; Igo et al. 2002). In this precleaved system the gRNA structure differs from the wild type in that its anchor is

modified to strongly bind to mRNA ABS, its guiding region is only a few nucleotides long, and the U-tail is replaced with a sequence complementary to the mRNA 5' fragment. In addition to bypassing the cleavage event, these modifications allow for less chimera formation, increasing overall editing efficiency (Carnes et al. 2005).

The enzyme cascade starts with a gRNA-directed endonuclease cleaving the mRNA that is base paired with a gRNA just 5' of the anchor duplex, at the 3' most ES or ES1. The cleavage generates a 5' fragment with a 3' hydroxyl and a 3' fragment with a 5' phosphate group. TUTase then inserts U(s) from free uridine triphosphate or exoUase deletes U(s) from the 5' end of the 3' fragment. Finally, a ligase re-joins both 5' and 3' fragments of the mRNA, completing a cycle, or "full-round", editing on a single editing site (Seiwert and Stuart 1994; Byrne et al. 1996; Cruz-Reyes and Sollner-Webb 1996; Kable et al. 1996; Seiwert et al. 1996; Simpson 2001; Simpson et al. 2003; Koslowsky 2004; Kang et al. 2005; Stuart et al. 2005; Carnes and Stuart 2007). The editosome proceeds through 1-30 editing sites per gRNA allowing formation of noncanonical G:U base pairing in addition to Watson-Crick pairing to zip up the gRNA with the mRNA (Blum et al. 1990; Stuart and Panigrahi 2002). The presence of TUTase and exoUase in distinct editosomes (Panigrahi et al. 2006; Carnes and Stuart 2007) suggested complex disassembly between distinct (insertion-deletion or deletion-insertion) sites. However, if the adjacent site is of the same type, the editosome is likely to work in a processive manner (Alatortsev et al. 2008). The cascade of reactions within a single site and the editosome switch between insertion/deletion sites implied dynamics and flexibility within each editosome and among the editosomes, in order to coordinate all the enzyme

activities and to recognize the different editing sites (Schnauffer et al. 2003; Koslowsky 2004; Lukes et al. 2005; Stuart et al. 2005).

As all the editing sites guided by a single gRNA are edited, the guiding region becomes fully complementary to the respective mRNA sequence, defining an editing block or domain (Kapushoc and Simpson 1999). The mRNA/gRNA complex is then separated by an RNA helicase mHel61p (named kinetoplastid RNA editing helicase KREH1) (Missel et al. 1997), the mRNA can be used for translation or hybridization with a new gRNA for editing of a next block, and the released gRNA can be recycled (Missel et al. 1999). Numerous editing blocks may exist within an mRNA, see figure 1.5D, and they can overlap depending on the transcript (Maslov and Simpson 1992; Corell et al. 1993).

Within an editing block, editing is thought to proceed in a 3' to 5' direction (Koslowsky et al. 1991). However, many partially edited intermediates have been identified containing edited sites upstream (5') of unedited sites (Koslowsky et al. 1991; Maslov and Simpson 1992; Sturm et al. 1992; Koslowsky and Yahampath 1997). One explanation is that editing is dictated by thermodynamic stabilities between the gRNA and mRNA, with less favorable sites being edited later (Koslowsky et al. 1991). Another explanation includes misediting by spurious gRNA binding, misguiding by the appropriate gRNA, or indiscriminate editing (Decker and Sollner-Webb 1990; Sturm and Simpson 1990; Sturm et al. 1992). No physiological consequences of aberrant editing have been found because the misedited transcripts usually did not complete editing at the 5' end, which is required to generate the AUG start translation codon (Koslowsky 2004). The recent discovery of alternative editing indicates that “misediting” can generate

distinct translation-competent mRNAs, increasing the mitochondrial gene diversity (Ochsenreiter and Hajduk 2006; Ochsenreiter et al. 2008). These studies detected translatable alternatively edited transcripts for the extensively edited A6, ND7, ND8, ND9 and COIII, and they also detected gRNA genes for the alternative edited regions. With still unknown functions, alternative editing may generate isoforms of a protein with different enzymatic activities and substrate specificities, different localization in the mitochondrion and altered ability to interact with other proteins or nucleic acids, like alternative RNA splicing.

### ***Other relevant RNA systems***

Our main interest focuses on early steps of kinetoplastid RNA editing, especially gRNA targeting of the mRNA and assembly of the gRNA/mRNA complex with the editosome. There are numerous systems involving RNA-RNA interaction and RNA processing in many organisms and some will be described below.

Due to discovery of large amounts of noncoding RNAs (ncRNAs) in cells, sometimes half of all transcripts, research into understanding RNA-RNA interactions gained importance in cellular metabolism and in biotechnological fields (Mattick and Makunin 2006). Noncoding RNA interaction partners can be proteins or other RNA molecules. Classic examples include rRNA and tRNA in mRNA translation, small nuclear RNAs (snRNAs) in splicing (McGrail and O'Keefe 2008), small nucleolar RNAs (snoRNAs) in modification of rRNA (Bachellerie et al. 2002), micro-RNAs from the RNAi pathway with their target mRNA (Perron and Provost 2008), ncRNAs from *E. coli* or sRNAs (Levine et al. 2007), and loop-loop interactions (Kolb et al. 2001; Paulus et al.

2004). Below are examples of systems involving and regulated by RNA-RNA interactions that may aid the understanding of the mRNA/gRNA interaction.

### **Antisense RNA**

Natural antisense RNAs are small noncoding transcripts (67-108 nt) of various structures that are complementary or partially complementary to longer target RNAs (Zeiler and Simons 1998). They can be engineered but exist naturally in bacteria, plants, mice, humans and certain protozoans (Militello et al. 2008). Upon binding to target RNA, they form stable complexes that can inhibit target function (e.g. in case of a target mRNA it will inhibit translation). The RNA-RNA hybridization is influenced by both RNA's primary, secondary and tertiary structures, thus linear-linear, linear-loop, or loop-loop interactions show dissimilar pathways (Eguchi et al. 1991). Linear-linear and linear-loop interactions directly evolve to stable complex formation whereas loop-loop interactions have a second step with intermediate state(s). Although the folding pathway may differ, natural antisense systems present common features, such as same limiting step, similar structural features, and equivalent reaction rates. The initial pairing is reversible, depends on the sequence, and occurs between short (5-7) complementary, single stranded sequences in either loops or unstructured sites. The presence of at least one unstructured single-stranded end is also common. Initiation occurs at similar second order association rates ( $10^5$ - $10^6$   $M^{-1}s^{-1}$ ) followed by rapid stable complex formation. Thus the rate of initial complex formation (not thermodynamic stability of that complex) is the limiting and determinant step for formation of a stable complex (Zeiler and Simons 1998).

In bacteria and certain plant and yeast cells, antisense RNAs regulate plasmid copy number. Plasmids are extra-chromosomal circular or linear double stranded DNA

carrying genes essential for survival. In the case of *E. coli* high-copy number plasmid ColE1, antisense RNAI (108 nt forming a 3 stem-loop structure with a 5' single-stranded tail) binds its target RNAII (100-360 nt) at the 5' end, which also contains 3 stem-loops. The pathway leading to a stable complex is complicated and may involve several intermediates. Nevertheless, the initial contact involves a single loop-loop interaction and the final RNAI/RNAII complex prevents RNAII from binding to ColE1 DNA, consequently inhibiting replication (Tomizawa 1986; Eguchi and Tomizawa 1991).

The low-copy number plasmid R1, also from *E. coli*, carries genes that confer resistance to antibiotics. R1 replicates only upon the binding of RepA protein, which is regulated by the antisense RNA CopA. Antisense CopA (90 nt) binds RepA mRNA at a site termed CopT (the leader region), inhibiting its translation. Both CopA and CopT have two stem-loops (I and II) separated by a single-stranded area. Pairing initiation starts with “kissing” of loops II, followed by rapid formation of an intermediate which undergoes first-order rearrangement towards the stable complex (Eguchi et al. 1991; Kolb et al. 2001). In ColE1, R1, and the majority of loop-loop antisense systems in prokaryotes, the antisense and target RNAs start the binding event through a “kissing” interaction, which is RNA contact limited to a few W-C base pairs formed by complementary sequences in the apical loops of two hairpins. This initial pairing significantly increases the stability of the loop-loop interaction promoting high second-order binding rate constants of approximately  $10^6 \text{ M}^{-1}\text{s}^{-1}$  (Craig et al. 1971; Eguchi and Tomizawa 1991). In these and many antisense mechanisms, efficiency and control depend on fast binding kinetics, where nucleation is the rate-limiting step. Nucleation is the formation of the first few base pairs in RNA-RNA hybridizations. This initial base

pairing is slow because it is not stabilized by stacking free energy; the second base pair can only form if the helix is aligned in a proper orientation. After the formation of approximately 3 bp the rest of the helix will close faster ( $10^6 \text{ s}^{-1}$ ) independent of base sequence. The helix dissociation rate depends on the length and number of base pairs of the helix (Craig et al. 1971).

There are a few well-characterized antisense regulated systems where stable complex formation initiates through a linear-linear RNA pairing (e.g. *hok/sok* post-segregational killing system of plasmid R1) or loop-linear pairing (e.g. RNA-IN/RNA-OUT of Tn10, controlling transposition frequency) (Eguchi et al. 1991; Zeiler and Simons 1998). Interestingly, if the linear 5' interacting end of RNA-IN is mutated to form a stem-loop structure, there will be some loop-loop interaction between RNA-IN and -OUT, but no stable complex formation (Zeiler and Simons 1998). The failure of complex formation for the loop-loop RNA-IN/OUT indicates that not all loop-loop interactions efficiently lead to stable complex formation. After the initial contact, one of the stems has to give way to enable single-strand passage of the complementary molecule and helix propagation. Thus, there are constraints for loop-loop and loop-linear interactions including nucleotide sequence within the loop, loop size, stem stability, and loops far from the ends of the molecules. Loops of 5-7 nt that contain a U-turn, or YUNR motif (where Y = pyrimidine, U = uracil, N = any nucleoside, R = purine) facilitate rapid binding. In this case, the phosphodiester backbone is hidden in the loop architecture promoting pairing by reducing the backbone repulsion (Franch and Gerdes 2000).

In addition to loops, bulges (unpaired bases in a helical stem) and helical junctions constitute important structural elements that can favor RNA-RNA interactions. Bulges lower the Gibbs free energy ( $\Delta G$ ) of the stem, promoting faster progression of a loop-loop interaction (Kolb et al. 2001) and RNA junctions promote flexibility in addition to stability and fast formation kinetics as it abolishes the necessity to form a complete duplex (Duckett et al. 1997; Kolb et al. 2001).

In linear-linear interactions (usually involving short oligoribonucleotides) the chain length and base composition strongly affect the stability of the complex, indicating that nearest neighbor interactions influence the contribution of each base pair to the stability of the helix (Eguchi et al. 1991; Xia et al. 1998).

Bi-molecular RNA interactions such as linear-linear, loop-linear, or loop-loop interactions are important in increasing the efficacy of antisense mechanisms. If the RNA molecules are in extended linear structures, then the specificity of the interaction is determined solely by the nucleotide sequence. Introducing structure into one or both interacting RNAs imposes an additional specificity that can affect the rate of binding and consequently increase the functional and genetic versatility of the interaction. Additionally, the formation of a specific structure within a group of RNAs (as we propose for the mRNA/gRNA complexes), allows any group member to interact with a group-specific protein, further increasing the specificity of the interaction (Eguchi et al. 1991; Zeiler and Simons 1998).

### **Micro-RNAs**

In contrast to antisense RNAs, micro-RNAs (miRNAs) are distinct examples of short 20-25 nt ncRNA transcripts that posttranscriptionally regulate gene expression in plants and

animals. The miRNA precursors (pri-miRNAs) are transcripts from intergenic or intron regions and are processed by RNase Drosha to a pre-miRNA hairpin and exported to the cytoplasm. Pre-miRNA is cleaved by Dicer [RNase III family, part of the RNA-induced silencing complex (RISC)], and becomes ds-miRNA. The miRNA dissociates or it is cleaved to generate a single stranded component that is loaded into RISC (contains Dicer and members of argonaute (Ago) proteins). Due to the existence of more than 15 Ago-associated proteins, there may be several RISC complexes working distinctly (MacRae et al. 2008). Nonetheless, the single stranded miRNA serves as a guide molecule containing the sequence for binding to specific mRNA sites, mainly on the 3'UTR (rarely at 5'UTR and ORF) consequently impeding translation or inducing deadenylation, cleavage or degradation (Lewis et al. 2005; Diederichs and Haber 2007). A single miRNA can have 100-200 target sites (Brennecke et al. 2005; Krek et al. 2005) and miRNAs are predicted to regulate 30% of the human genome (Lewis et al. 2005). There are at least 6 features observed in miRNA-target RNA interactions. 1) Most miRNA sites are canonical, meaning they involve full pairing. 2) In many cases, pairing at the 5' end of the miRNA (the "seed" site) is sufficient (minimum of 7 complementary and continuous base pairing). If seed binding is strong, it rarely needs pairing from its 3' end. 3) If the 5' end seed pairing is weak it can still be functional only if the 3' end promotes strong compensatory pairing. 4) The compensatory 3' end pairing can be 4-6 continuous base pairs, or 7-8 bp that contain G:U, single nucleotide bulges, or mismatches. 5) A few G:U base pairs, single-nucleotide bulges or mismatches are allowed in the 5' seed sequence if the 3' end pairing of the miRNA is compensatory. 6) The position of the seed site affects functionality but there is no correlation with free energy (Brennecke et al. 2005; Didiano

and Hobert 2006). Although the above conclusions exist in the literature they were not tested using all existent miRNAs, thus should be used with precaution (Didiano and Hobert 2006).

The branch of the interference RNA (RNAi) pathway where the miRNA is loaded into RISC and binds to a specific sequence within the mRNA resembles the initial steps in RNA editing. The 5' miRNA seed sequence is similar in size and features to the gRNA anchor region. Both RNA-RNA interactions allow G:U base pairing and it seems that the stronger the 5' pairing the less necessary the pairing at the 3' region will be. The lack of consensus sequence, the potential existence of distinct multiprotein complexes, and unknown accessory proteins (Preall and Sontheimer 2005) are other features shared by both systems.

### **Other RNA-RNA interactions including RNA editing in other systems**

There is a great number of RNA-catalyzed and protein-catalyzed RNA processing mechanisms that involve RNA-RNA interactions and remarkable site-specific modification. It is beyond the scope of this dissertation to review all of them; instead, I will talk about some systems while citing a few reviews to highlight how important RNA-RNA interactions and RNA structure can be in catalytic processes.

Among the RNA-catalyzed machines, a few may require protein factors to work in vivo, like some Group I and II RNA self-splicing introns and spliceosome mediated mRNA splicing. While RNA structural motifs are fundamental and the only components of self-cleaving ribozymes (hammerhead, hairpin, hepatitis delta, neurospora VS), in RNase P catalysis, a protein subunit plays the structural role. RNase P is a RNP where

the RNA subunit contains the active site responsible for 5' processing of precursor tRNAs (Walter and Engelke 2002; Hoogstraten and Sumita 2007).

The protein-catalyzed RNA systems include the previously mentioned editosome and RISC, telomerase and snoRNPs (Walter and Engelke 2002; Hoogstraten and Sumita 2007). Small nucleolar RNAs (snoRNAs) are similar to trypanosome gRNAs in that they associate with catalytic proteins and are responsible for recognition of target RNA. The formation of 2'O-methylated nucleosides (Nm) and the conversion of uridine to pseudouridine ( $\Psi$ ) are two common snoRNA-guided modifications among rRNAs, splicing snRNAs, and tRNAs, mediated by large heterogeneous snoRNPs (Decatur and Fournier 2003). The Nm guide snoRNA contains one or two pairs of specific sequence elements called boxes C and D, and C' and D'. These boxes are required for methylation as they are spatial determinants and affect protein binding. Like in the Nm case, the  $\Psi$  guide snoRNA also has specific sequences, the boxes H and ACA that are required for the processing event, protein binding, and localization (Cavaille and Bachellerie 1998; Kiss 2001). In addition to directing site-specific modifications, snoRNAs can induce rRNA cleavage by pairing to rRNA and presenting the cleavage site to the endoribonuclease, a function similar to the trypanosome gRNA (Stoltzfus 1999).

RNA editing in other organisms does not necessarily involve bimolecular RNA interactions, yet aid in the study of trypanosome editing because site-specificity depends on RNA structure, cofactors, and enzymes as catalysts. The deamination of cytidine to uridine (C $\rightarrow$ U) by the cytidine deaminase APOBEC1 in the mRNA encoding apolipoprotein B, and deamination of adenosine to inosine (A $\rightarrow$ I) by adenosine deaminases that act on RNA (ADAR) in mRNAs encoding glutamate-sensitive ion

channels and serotonin receptors, are two well known forms of RNA editing by base modification in mammals (Gott and Emeson 2000). ADAR and APOBEC1 do not use an RNA molecule to guide their modifications. Instead, RNA sequence and mainly RNA structural features, such as double strand sites, mismatches, bulges, and loops, are fundamental for target recognition and site-specificity (Connell and Simpson 1998; Richardson et al. 1998; Gott and Emeson 2000; Bass 2002; Koslowsky 2004).

RNA editing in plant organelles is another example devoid of RNA-RNA interactions but is similar to trypanosomes in that it involves RNA modification in-organelle and comprises changes at hundreds of sites. In both plant mitochondria and chloroplasts, editing is site-specific, does not need guide RNA molecules and the modifications are conversions of C→U, or very rarely U→C (Brennicke et al. 1999). The determinants of target recognition are not fully known, but there is a preferred sequence around the editing site consisting of two pyrimidines (Y) upstream and a purine (R) downstream of the edited nucleotide (YYCR) (Mulligan et al. 2007). In addition to the preferred sequence, cis elements of unspecified sequence, from -20 to +10 of the editing site, are required for RNA editing in these organelles. In plastids, these cis elements may be the binding site for 450 pentatricopeptide repeat proteins (PPR). PPRs are RNA-binding proteins that function in post-transcription processes (including RNA splicing, RNA editing, RNA cleavage and translation). With so many of these proteins in the chloroplast, a distinct PPR protein is proposed to exist to regulate each individual editing site (Shikanai 2006).

Each of the above examples adds to the understanding of RNA editing in trypanosomes. Although the players and mechanisms differ, the importance of RNA

structure for proper function is common to all (Sommer et al. 1991; Connell and Simpson 1998; Koslowsky 2004; Glazov et al. 2006; Lunde et al. 2007; Wakeman et al. 2007).

### **Concluding Remarks**

Despite all the progress in identifying the components of the editing machinery, there are many unanswered questions concerning the temporal association of gRNAs, mRNAs and the editing proteins into a ribonucleoprotein (RNP) complex.

Nascent RNAs are likely to be bound by proteins (Lunde et al. 2007) thus, it has been suggested that mRNAs and gRNAs may be initially associated with mitochondrial proteins and then brought together to form the anchor duplex within the editosome (Madison-Antenucci et al. 2002). Some gRNA binding proteins are thought to influence mRNA editing in a transcript-specific manner (Goringer et al. 1994; Koller et al. 1994; Peris et al. 1994; Read et al. 1994; Bringaud et al. 1995; Byrne et al. 1995; Leegwater et al. 1995; Shu et al. 1995; Hayman and Read 1999; Madison-Antenucci and Hajduk 2001; Zikova et al. 2006). For instance, RBP16 was shown to bind preferentially to the gRNA U-tail and stimulate mainly CYb editing (Pelletier and Read 2003; Miller et al. 2006). The MRP complex [MRP1 (gBP21) / MRP2 (gBP25)] recognizes common structural elements within the gRNA anchor and guiding region, facilitating gRNA/mRNA annealing by keeping the anchor region single stranded (Koller et al. 1997; Lambert et al. 1999; Muller et al. 2001; Muller and Goringer 2002; Vondruskova et al. 2005; Schumacher et al. 2006). On other hand, proteins involved in mRNA binding, including REAP-1 (Madison-Antenucci and Hajduk 2001; Hans et al. 2007), RBP16, and others (Halbig et al. 2006) have also been found but with yet unspecified function in editing.

Alternatively, gRNAs and mRNAs may be individually loaded into the editosome prior to hybridization. mRNA editing domains fold into stem-loops that could be target sites for the editing endoribonuclease, which would then recruit a gRNA to proceed with editing (Piller et al. 1995). gRNAs alone (no mRNAs) could assemble with the editosome, as they have been isolated in conjunction with the editosome or a RNP (Pollard et al. 1992; Shu et al. 1995). KREPA6 is a structural protein existent in all editosomes, found to bind stretches of Us within RNAs. As there are more than 1200 gRNAs, KREPA6 could indiscriminately bind any gRNA U-tail further facilitating hybridization with the cognate mRNA (Tarun et al. 2008). Another conclusion taken from the fact that gRNAs were solely detected in the editosome was that gRNAs must share a common information that places them in a group that is recognized by the same protein(s) (Goringer et al. 1994). There is not a consensus sequence among the gRNAs besides the U-tail, but the secondary structure of at least four of them has similarities. Through enzymatic and chemical probing it was found that four gRNAs shared a low-stability, dual stem-loop structure (Schmid et al. 1995; Schmid et al. 1996; Fig. 2.6A). Even without consensus structures, the 3-D model obtained later for one gRNA pointed to the importance of tertiary interactions stabilizing the molecule (Hermann et al. 1997). More recently, results from the solved MRP1/MRP2-gRNA crystal structure corroborate previous findings where the MRP complex binds and unfolds the 5' most gRNA stem-loop via nonspecific contacts, possibly facilitating further hybridization with mRNA (Muller et al. 2001; Muller and Goringer 2002; Aphasizhev et al. 2003; Schumacher et al. 2006).

mRNA/gRNA complexes may properly form spontaneously in the absence of proteins, and the complex itself would facilitate the recruitment and assembly of the

editing machinery (Rusche et al. 1997; Leung and Koslowsky 2001a; Yu and Koslowsky 2006). Structural elements of the interacting mRNA/gRNAs are critical for editing site recognition and editing efficiency (Cruz-Reyes and Sollner-Webb 1996; Seiwert et al. 1996; Leung and Koslowsky 1999; Lawson et al. 2001; Leung and Koslowsky 2001a; 2001b; Koslowsky 2004; Golden and Hajduk 2006; Yu and Koslowsky 2006). For instance, the presence of editing sites within a single-stranded loop is important for cleavage by the kinetoplastid endonuclease (Harris et al. 1992; Simpson et al. 1992). The site specificity of the kinetoplastid endonuclease resembles that of some well-studied RNase III, like the *E. coli* or the RNase III of the miRNA pathway in vertebrates. For these RNase III, RNA structure like single-stranded loops, including the size of the loops are also determinants of cleavage specificity (Robertson 1982; Zhang and Nicholson 1997; Nicholson 1999; Ritchie et al. 2007). As opposed to loops, Watson-Crick (W-C) base pairs negatively control RNase III cleavage by inhibiting enzyme binding (Zhang and Nicholson 1997). In kinetoplastid editing, site recognition by the endoribonuclease is still unclear and very complicated as there are thousands of distinct editing sites to insert 3583 and delete 322 Us within the 12 mRNAs. A collection of studies have verified that structural features of the complex mRNA/gRNA influence editing (Piller et al. 1995; Byrne et al. 1996; Schmid et al. 1996; Seiwert et al. 1996; Cruz-Reyes et al. 1998; Burgess et al. 1999; Leung and Koslowsky 1999; Burgess and Stuart 2000; Igo et al. 2000; Cruz-Reyes et al. 2001; Lawson et al. 2001; Leung and Koslowsky 2001a; 2001b; Igo et al. 2002; Oppegard et al. 2003; Pai et al. 2003; Koslowsky 2004; Golden and Hajduk 2006; Alatorsev et al. 2008). From these studies, we can summarize the following. Anchor duplex formation and length is fundamental for editing in vitro.

Some base-pairing between gRNA and its cognate mRNA upstream of the editing site is not required but enhances editing efficiency. This upstream stability is greatly influenced by the presence of the gRNA U-tail. Although sequence may influence ES recognition, sequence alone seems unlikely to be sufficient for the differential recognition by the three distinct *T. brucei* endoribonucleases, especially due to the existence of hundreds of editing sites. Moreover, no conservative sequence elements have been found around the ESs or in the gRNA (Stuart et al. 2005).

Independently of how mRNA/gRNAs are brought together or assembled into the editosome, in vitro editing assays have been successfully established without the need of accessory proteins (Seiwert and Stuart 1994; Cruz-Reyes and Sollner-Webb 1996; Kable et al. 1996; Seiwert et al. 1996; Kang et al. 2005; Carnes and Stuart 2007).

### ***Focus of this Dissertation***

Our interest is in mRNA/gRNA structure and also gRNA targeting. More specifically, I am interested in the structure of the ATPase subunit 6 mRNAs alone and in complex during distinct stages of editing, as well as the forces and thermodynamics driving the interaction.

We hypothesize that mRNAs and cognate gRNAs can hybridize in the absence of proteins and that formation of a gRNA/mRNA anchor duplex initiates editing (Rusche et al. 1997; Stuart et al. 2005). Upon anchor helix pairing, complex re-organization may be necessary to allow formation of U-tail helix and subsequent tertiary interactions, further increasing complex stability and likely presenting the ES to the endoribonuclease (Yu and Koslowsky 2006). The pathway for complex formation is likely to be distinct between mRNA/gRNA pairs due to their distinct primary and secondary sequences. For

instance, gA6-14 binds its cognate mRNA with roughly 100 times greater affinity ( $K_D$  in nM range) when compared with gCYb-558 and its mRNA ( $K_D$  in  $\mu$ M range) (Koslowsky 2004). The A6 mRNA secondary structure has the ABS within a mostly single stranded area as opposed to the CYb ABS involved in a stable stem-loop. The A6 stronger binding affinity is due to faster association and slower dissociation rate constants, as measured by surface plasmon resonance (Yu 2006). Additionally, the above A6 pair is used as a standard, as it is the only pair that undergoes a full round of editing in vitro without the need of accessory proteins or modifications (Seiwert and Stuart 1994). The CYb pair will undergo editing in vitro only with help from RBP16 and mutations that increase the complex stability (Miller et al. 2006).

Previous studies in our laboratory found no consensus structure among distinct mRNA/gRNA pairs when analyzed by a computer algorithm (Leung and Koslowsky 1999). Only after crosslinking the gRNA to the mRNA and incorporating constraints into the same computer modeling program, were all three pairs shown to fold into a three-helical structure (Leung and Koslowsky 1999) that was similar to those predicted previously (Blum and Simpson 1990). Through crosslinking and structure probing studies, this predicted secondary structure was later confirmed to exist in solution for the unedited and partially edited CYb/gCYb-558 pair (Leung and Koslowsky 2001a; Yu and Koslowsky 2006). In brief, CYb/gCYb-558 presented an anchor helix (or 3' helix, between the gRNA anchor and the mRNA anchor binding site), a second helix formed within the gRNA guiding region itself, and the third helix (U-tail or 5' helix) involved the purine rich region of the mRNA, just upstream of the editing site, and the gRNA U-tail (Fig. 1.7). Interestingly, structure probing of mRNAs that were partially edited through

three sites indicated that the U-tail helix and gRNA stem-loop can undergo structural arrangements that maintain the core 3-helical structure (Yu and Koslowsky 2006), which suggests that this conformation is relevant for editing.

Akin to antisense RNAs, mRNA-gRNA hybridization may present distinct folding pathways and limiting steps; but we hypothesize that all complexes fold into a common structure. The next chapter of this thesis describes the work conducted to determine the structure in solution for the A6/gA6-14 pair, unedited and partially edited, and how it resembles the previously published CYb/gCYb-558 complex. In chapter three, I investigated how mRNA structure alone can influence mRNA/gRNA complex formation. Chapter 4 describes the study conducted to determine the thermodynamic parameters for the A6/gA6-14 pair and how the gRNA U-tail affects the interaction. Finally, Chapter 5 includes the concluding remarks for the project.

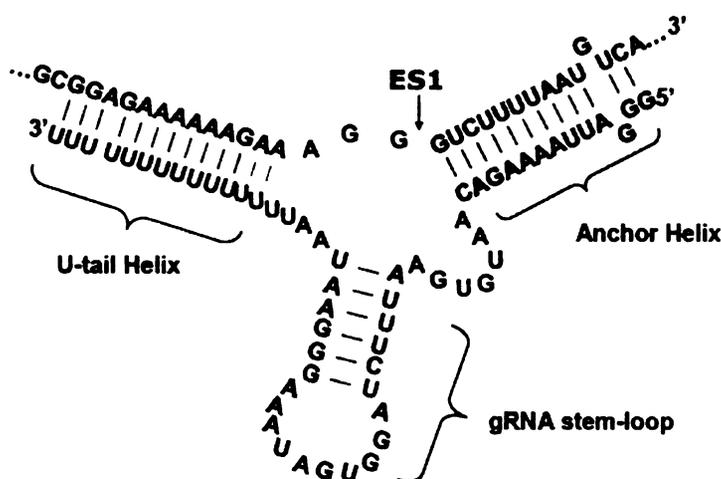


Figure 1.7 Secondary structure model for the mRNA/gRNA complex. The three helices, anchor helix (or 3' helix), gRNA stem-loop, and U-tail helix (or 5' helix) expose the editing site ES1 just 5' of the anchor helix. Modified from Yu LE. 2006. Elucidating the structure and kinetics of the apocytochrome B mRNA/gRNA complex in *Trypanosoma brucei* mitochondria. Cell and Molecular Biology. East Lansing: Michigan State University. pp 205.

## References

- Adler, B.K., and Hajduk, S.L. 1997. Guide RNA requirement for editing-site-specific endonucleolytic cleavage of preedited mRNA by mitochondrial ribonucleoprotein particles in *Trypanosoma brucei*. *Mol Cell Biol* **17**: 5377-5385.
- Alatortsev, V.S., Cruz-Reyes, J., Zhelonkina, A.G., and Sollner-Webb, B. 2008. *Trypanosoma brucei* RNA editing: coupled cycles of U-deletion reveal processive activity of the editing complex. *Mol Cell Biol*.
- Aphasizhev, R., Aphasizheva, I., Nelson, R.E., Gao, G., Simpson, A.M., Kang, X., Falick, A.M., Sbicego, S., and Simpson, L. 2003. Isolation of a U-insertion/deletion editing complex from *Leishmania tarentolae* mitochondria. *Embo J* **22**: 913-924.
- Aphasizhev, R., Aphasizheva, I., Nelson, R.E., and Simpson, L. 2003. A 100-kD complex of two RNA-binding proteins from mitochondria of *Leishmania tarentolae* catalyzes RNA annealing and interacts with several RNA editing components. *RNA* **9**: 62-76.
- Aphasizhev, R., Aphasizheva, I., and Simpson, L. 2003. A tale of two TUTases. *Proc Natl Acad Sci U S A* **100**: 10617-10622.
- Aphasizhev, R., Sbicego, S., Peris, M., Jang, S.H., Aphasizheva, I., Simpson, A.M., Rivlin, A., and Simpson, L. 2002. Trypanosome mitochondrial 3' terminal uridylyl transferase (TUTase): the key enzyme in U-insertion/deletion RNA editing. *Cell* **108**: 637-648.
- Babbarwal, V.K., Fleck, M., Ernst, N.L., Schnauffer, A., and Stuart, K. 2007. An essential role of KREPB4 in RNA editing and structural integrity of the editosome in *Trypanosoma brucei*. *RNA* **13**: 737-744.
- Bachelierie, J.P., Cavaille, J., and Huttenhofer, A. 2002. The expanding snoRNA world. *Biochimie* **84**: 775-790.
- Barrett, M.P., Burchmore, R.J., Stich, A., Lazzari, J.O., Frasch, A.C., Cazzulo, J.J., and Krishna, S. 2003. The trypanosomiases. *Lancet* **362**: 1469-1480.
- Bass, B.L. 2002. RNA editing by adenosine deaminases that act on RNA. *Annu Rev Biochem* **71**: 817-846.
- Benne, R. 1994. RNA editing in trypanosomes. *Eur J Biochem* **221**: 9-23.
- Benne, R., De Vries, B.F., Van den Burg, J., and Klaver, B. 1983. The nucleotide sequence of a segment of *Trypanosoma brucei* mitochondrial maxi-circle DNA that contains the gene for apocytochrome b and some unusual unassigned reading frames. *Nucleic Acids Res* **11**: 6925-6941.

- Benne, R., Van den Burg, J., Brakenhoff, J.P., Sloof, P., Van Boom, J.H., and Tromp, M.C. 1986. Major transcript of the frameshifted coxII gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA. *Cell* **46**: 819-826.
- Bern, C., Montgomery, S.P., Herwaldt, B.L., Rassi, A., Jr., Marin-Neto, J.A., Dantas, R.O., Maguire, J.H., Acquatella, H., Morillo, C., Kirchhoff, L.V., et al. 2007. Evaluation and treatment of chagas disease in the United States: a systematic review. *Jama* **298**: 2171-2181.
- Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renault, H., Bartholomeu, D.C., Lennard, N.J., Caler, E., Hamlin, N.E., Haas, B., et al. 2005. The genome of the African trypanosome *Trypanosoma brucei*. *Science* **309**: 416-422.
- Besteiro, S., Barrett, M.P., Riviere, L., and Bringaud, F. 2005. Energy generation in insect stages of *Trypanosoma brucei*: metabolism in flux. *Trends Parasitol* **21**: 185-191.
- Beverley, S.M. 2003. Protozomics: trypanosomatid parasite genetics comes of age. *Nat Rev Genet* **4**: 11-19.
- Bhat, G.J., Koslowsky, D.J., Feagin, J.E., Smiley, B.L., and Stuart, K. 1990. An extensively edited mitochondrial transcript in kinetoplasts encodes a protein homologous to ATPase subunit 6. *Cell* **61**: 885-894.
- Blum, B., Bakalara, N., and Simpson, L. 1990. A model for RNA editing in kinetoplastid mitochondria: "guide" RNA molecules transcribed from maxicircle DNA provide the edited information. *Cell* **60**: 189-198.
- Blum, B., and Simpson, L. 1990. Guide RNAs in kinetoplastid mitochondria have a nonencoded 3' oligo(U) tail involved in recognition of the preedited region. *Cell* **62**: 391-397.
- Blum, B., Sturm, N.R., Simpson, A.M., and Simpson, L. 1991. Chimeric gRNA-mRNA molecules with oligo(U) tails covalently linked at sites of RNA editing suggest that U addition occurs by transesterification. *Cell* **65**: 543-550.
- Brecht, M., Niemann, M., Schluter, E., Muller, U.F., Stuart, K., and Goring, H.U. 2005. TbMP42, a protein component of the RNA editing complex in African trypanosomes, has endo-exoribonuclease activity. *Mol Cell* **17**: 621-630.
- Brennecke, J., Stark, A., Russell, R.B., and Cohen, S.M. 2005. Principles of microRNA-target recognition. *PLoS Biol* **3**: e85.
- Brennicke, A., Marchfelder, A., and Binder, S. 1999. RNA editing. *FEMS Microbiol Rev* **23**: 297-316.

- Bringaud, F., Peris, M., Zen, K.H., and Simpson, L. 1995. Characterization of two nuclear-encoded protein components of mitochondrial ribonucleoprotein complexes from *Leishmania tarentolae*. *Mol Biochem Parasitol* **71**: 65-79.
- Brun, R., Hecker, H., and Lun, Z.R. 1998. *Trypanosoma evansi* and *T. equiperdum*: distribution, biology, treatment and phylogenetic relationship (a review). *Vet Parasitol* **79**: 95-107.
- Burgess, M.L., Heidmann, S., and Stuart, K. 1999. Kinetoplastid RNA editing does not require the terminal 3' hydroxyl of guide RNA, but modifications to the guide RNA terminus can inhibit in vitro U insertion. *RNA* **5**: 883-892.
- Burgess, M.L., and Stuart, K. 2000. Sequence bias in edited kinetoplastid RNAs. *RNA* **6**: 1492-1497.
- Buscher, P. 2002. Diagnosis of human and animal african trypanosomiasis. In *The African Trypanosomes*. (eds. S. Black, and J. Seed), pp. 192. Springer, New York.
- Byrne, E., Bringaud, F., and Simpson, L. 1995. RNA-protein interactions in the ribonucleoprotein T-complexes in a mitochondrial extract from *Leishmania tarentolae*. *Mol Biochem Parasitol* **72**: 65-76.
- Byrne, E.M., Connell, G.J., and Simpson, L. 1996. Guide RNA-directed uridine insertion RNA editing in vitro. *Embo J* **15**: 6758-6765.
- Camargo, E.P. 1999. *Phytomonas* and other trypanosomatid parasites of plants and fruit. *Adv Parasitol* **42**: 29-112.
- Carnes, J., and Stuart, K.D. 2007. Uridine insertion/deletion editing activities. *Methods Enzymol* **424**: 25-54.
- Carnes, J., Trotter, J.R., Ernst, N.L., Steinberg, A., and Stuart, K. 2005. An essential RNase III insertion editing endonuclease in *Trypanosoma brucei*. *Proc Natl Acad Sci USA* **102**: 16614-16619.
- Carnes, J., Trotter, J.R., Peltan, A., Fleck, M., and Stuart, K. 2008. RNA editing in *Trypanosoma brucei* requires three different editosomes. *Mol Cell Biol* **28**: 122-130.
- Cattand, P., Jannin, J., and Lucas, P. 2001. Sleeping sickness surveillance: an essential step towards elimination. *Trop Med Int Health* **6**: 348-361.
- Cavaille, J., and Bachellerie, J.P. 1998. SnoRNA-guided ribose methylation of rRNA: structural features of the guide RNA duplex influencing the extent of the reaction. *Nucleic Acids Res* **26**: 1576-1587.
- Chappuis, F., Loutan, L., Simarro, P., Lejon, V., and Buscher, P. 2005. Options for field diagnosis of human african trypanosomiasis. *Clin Microbiol Rev* **18**: 133-146.

- Clement, S.L., Mingler, M.K., and Koslowsky, D.J. 2004. An intragenic guide RNA location suggests a complex mechanism for mitochondrial gene expression in *Trypanosoma brucei*. *Eukaryot Cell* **3**: 862-869.
- Connell, G.J., and Simpson, L. 1998. Role of RNA Structure in RNA Editing. In *RNA Structure and Function*. (eds. R.W. Simons, and M. Grunberg-Manago), pp. 641-667. Cold Spring Harbor Laboratory Press, Plainview, NY.
- Corell, R.A., Feagin, J.E., Riley, G.R., Strickland, T., Guderian, J.A., Myler, P.J., and Stuart, K. 1993. *Trypanosoma brucei* minicircles encode multiple guide RNAs which can direct editing of extensively overlapping sequences. *Nucleic Acids Res* **21**: 4313-4320.
- Coustou, V., Besteiro, S., Biran, M., Diolez, P., Bouchaud, V., Voisin, P., Michels, P.A., Canioni, P., Baltz, T., and Bringaud, F. 2003. ATP generation in the *Trypanosoma brucei* procyclic form: cytosolic substrate level is essential, but not oxidative phosphorylation. *J Biol Chem* **278**: 49625-49635.
- Craig, M.E., Crothers, D.M., and Doty, P. 1971. Relaxation kinetics of dimer formation by self complementary oligonucleotides. *J Mol Biol* **62**: 383-401.
- Cruz-Reyes, J., Rusche, L.N., Piller, K.J., and Sollner-Webb, B. 1998. *T. brucei* RNA editing: adenosine nucleotides inversely affect U-deletion and U-insertion reactions at mRNA cleavage. *Mol Cell* **1**: 401-409.
- Cruz-Reyes, J., Rusche, L.N., and Sollner-Webb, B. 1998. *Trypanosoma brucei* U insertion and U deletion activities co-purify with an enzymatic editing complex but are differentially optimized. *Nucleic Acids Res* **26**: 3634-3639.
- Cruz-Reyes, J., and Sollner-Webb, B. 1996. Trypanosome U-deletional RNA editing involves guide RNA-directed endonuclease cleavage, terminal U exonuclease, and RNA ligase activities. *Proc Natl Acad Sci USA* **93**: 8901-8906.
- Cruz-Reyes, J., Zhelonkina, A., Rusche, L., and Sollner-Webb, B. 2001. Trypanosome RNA editing: simple guide RNA features enhance U deletion 100-fold. *Mol Cell Biol* **21**: 884-892.
- Cruz-Reyes, J., Zhelonkina, A.G., Huang, C.E., and Sollner-Webb, B. 2002. Distinct functions of two RNA ligases in active *Trypanosoma brucei* RNA editing complexes. *Mol Cell Biol* **22**: 4652-4660.
- Cunningham, I., and Honigberg, B.M. 1977. Infectivity reacquisition by *Trypanosoma brucei brucei* cultivated with tsetse salivary glands. *Science* **197**: 1279-1282.
- Deborggraeve, S., Claes, F., Laurent, T., Mertens, P., Leclipteux, T., Dujardin, J.C., Herdewijn, P., and Buscher, P. 2006. Molecular dipstick test for diagnosis of sleeping sickness. *J Clin Microbiol* **44**: 2884-2889.

- Decatur, W.A., and Fournier, M.J. 2003. RNA-guided nucleotide modification of ribosomal and other RNAs. *J Biol Chem* **278**: 695-698.
- Decker, C.J., and Sollner-Webb, B. 1990. RNA editing involves indiscriminate U changes throughout precisely defined editing domains. *Cell* **61**: 1001-1011.
- Delespaux, V., and de Koning, H.P. 2007. Drugs and drug resistance in African trypanosomiasis. *Drug Resist Updat* **10**: 30-50.
- Didiano, D., and Hobert, O. 2006. Perfect seed pairing is not a generally reliable predictor for miRNA-target interactions. *Nat Struct Mol Biol* **13**: 849-851.
- Diederichs, S., and Haber, D.A. 2007. Dual role for argonautes in microRNA processing and posttranscriptional regulation of microRNA expression. *Cell* **131**: 1097-1108.
- Docampo, R., and Moreno, S.N. 2003. Current chemotherapy of human African trypanosomiasis. *Parasitol Res* **90 Supp 1**: S10-13.
- Domingo, G.J., Palazzo, S.S., Wang, B., Pannicucci, B., Salavati, R., and Stuart, K.D. 2003. Dyskinetoplastic Trypanosoma brucei contains functional editing complexes. *Eukaryot Cell* **2**: 569-577.
- Duckett, D.R., Murchie, A.I., Clegg, R.M., Bassi, G.S., Giraud-Panis, M.J., and Lilley, D.M. 1997. Nucleic acid structure and recognition. *Biophys Chem* **68**: 53-62.
- Eguchi, Y., Itoh, T., and Tomizawa, J. 1991. Antisense RNA. *Annu Rev Biochem* **60**: 631-652.
- Eguchi, Y., and Tomizawa, J. 1991. Complexes formed by complementary RNA stem-loops. Their formations, structures and interaction with ColE1 Rom protein. *J Mol Biol* **220**: 831-842.
- Ernst, N.L., Pannicucci, B., Igo, R.P., Jr., Panigrahi, A.K., Salavati, R., and Stuart, K. 2003. TbMP57 is a 3' terminal uridylyl transferase (TUTase) of the Trypanosoma brucei editosome. *Mol Cell* **11**: 1525-1536.
- Etheridge, R.D., Aphasizheva, I., Gershon, P.D., and Aphasizhev, R. 2008. 3' adenylation determines mRNA abundance and monitors completion of RNA editing in T. brucei mitochondria. *Embo J*.
- Feagin, J.E., Abraham, J.M., and Stuart, K. 1988. Extensive editing of the cytochrome c oxidase III transcript in Trypanosoma brucei. *Cell* **53**: 413-422.
- Feagin, J.E., Jasmer, D.P., and Stuart, K. 1987. Developmentally regulated addition of nucleotides within apocytochrome b transcripts in Trypanosoma brucei. *Cell* **49**: 337-345.

- Fevre, E.M., Odiit, M., Coleman, P.G., Woolhouse, M.E., and Welburn, S.C. 2008. Estimating the burden of rhodesiense sleeping sickness during an outbreak in Serere, eastern Uganda. *BMC Public Health* **8**: 96.
- Franch, T., and Gerdes, K. 2000. U-turns and regulatory RNAs. *Curr Opin Microbiol* **3**: 159-164.
- Gibson, W. 2007. Resolution of the species problem in African trypanosomes. *Int J Parasitol* **37**: 829-838.
- Gibson, W., Peacock, L., Ferris, V., Williams, K., and Bailey, M. 2008. The use of yellow fluorescent hybrids to indicate mating in *Trypanosoma brucei*. *Parasit Vectors* **1**: 4.
- Glazov, E.A., Pheasant, M., Nahkuri, S., and Mattick, J.S. 2006. Evidence for control of splicing by alternative RNA secondary structures in Dipteran homothorax pre-mRNA. *RNA Biol* **3**: 36-39.
- Golden, D.E., and Hajduk, S.L. 2006. The importance of RNA structure in RNA editing and a potential proofreading mechanism for correct guide RNA:pre-mRNA binary complex formation. *J Mol Biol* **359**: 585-596.
- Goring, H.U., Koslowsky, D.J., Morales, T.H., and Stuart, K. 1994. The formation of mitochondrial ribonucleoprotein complexes involving guide RNA molecules in *Trypanosoma brucei*. *Proc Natl Acad Sci USA* **91**: 1776-1780.
- Gott, J.M., and Emeson, R.B. 2000. Functions and mechanisms of RNA editing. *Annu Rev Genet* **34**: 499-531.
- Grosjean, N.L., Vrable, R.A., Murphy, A.J., and Mansfield, L.S. 2003. Seroprevalence of antibodies against *Leishmania* spp among dogs in the United States. *J Am Vet Med Assoc* **222**: 603-606.
- Hajduk, S., and Sabatini, R.S. 1998. Mitochondrial mRNA editing in kinetoplastid protozoa. In *Modification and editing of RNA*. (eds. H. Grosjean, and R. Benne). ASM Press, Washington, DC.
- Halbig, K., Sacharidou, A., De Nova-Ocampo, M., and Cruz-Reyes, J. 2006. Preferential interaction of a 25kDa protein with an A6 pre-mRNA substrate for RNA editing in *Trypanosoma brucei*. *Int J Parasitol* **36**: 1295-1304.
- Hannaert, V., Bringaud, F., Opperdoes, F.R., and Michels, P.A. 2003. Evolution of energy metabolism and its compartmentation in Kinetoplastida. *Kinetoplastid Biol Dis* **2**: 11.
- Hans, J., Hajduk, S.L., and Madison-Antenucci, S. 2007. RNA-editing-associated protein 1 null mutant reveals link to mitochondrial RNA stability. *Rna* **13**: 881-889.

- Harris, M., Decker, C., Sollner-Webb, B., and Hajduk, S. 1992. Specific cleavage of pre-edited mRNAs in trypanosome mitochondrial extracts. *Mol Cell Biol* **12**: 2591-2598.
- Hashimi, H., Zikova, A., Panigrahi, A.K., Stuart, K.D., and Lukes, J. 2008. TbRGG1, an essential protein involved in kinetoplastid RNA metabolism that is associated with a novel multiprotein complex. *Rna* **14**: 970-980.
- Hayman, M.L., and Read, L.K. 1999. Trypanosoma brucei RBP16 is a mitochondrial Y-box family protein with guide RNA binding activity. *J Biol Chem* **274**: 12067-12074.
- Hellemond, J.J., Bakker, B.M., and Tielens, A.G. 2005. Energy metabolism and its compartmentation in Trypanosoma brucei. *Adv Microb Physiol* **50**: 199-226.
- Hermann, T., Schmid, B., Heumann, H., and Goring, H.U. 1997. A three-dimensional working model for a guide RNA from Trypanosoma brucei. *Nucleic Acids Res* **25**: 2311-2318.
- Hirumi, H., and Hirumi, K. 1989. Continuous cultivation of Trypanosoma brucei blood stream forms in a medium containing a low concentration of serum protein without feeder cell layers. *J Parasitol* **75**: 985-989.
- Hoogstraten, C.G., and Sumita, M. 2007. Structure-function relationships in RNA and RNP enzymes: recent advances. *Biopolymers* **87**: 317-328.
- Huang, C.E., Cruz-Reyes, J., Zhelonkina, A.G., O'Hearn, S., Wirtz, E., and Sollner-Webb, B. 2001. Roles for ligases in the RNA editing complex of Trypanosoma brucei: band IV is needed for U-deletion and RNA repair. *Embo J* **20**: 4694-4703.
- Igo, R.P., Jr., Lawson, S.D., and Stuart, K. 2002. RNA sequence and base pairing effects on insertion editing in Trypanosoma brucei. *Mol Cell Biol* **22**: 1567-1576.
- Igo, R.P., Jr., Palazzo, S.S., Burgess, M.L., Panigrahi, A.K., and Stuart, K. 2000. Uridylate addition and RNA ligation contribute to the specificity of kinetoplastid insertion RNA editing. *Mol Cell Biol* **20**: 8447-8457.
- Igo, R.P., Jr., Weston, D.S., Ernst, N.L., Panigrahi, A.K., Salavati, R., and Stuart, K. 2002. Role of uridylate-specific exoribonuclease activity in Trypanosoma brucei RNA editing. *Eukaryot Cell* **1**: 112-118.
- Jannin, J., and Cattand, P. 2004. Treatment and control of human African trypanosomiasis. *Curr Opin Infect Dis* **17**: 565-571.
- Kable, M.L., Seiwert, S.D., Heidmann, S., and Stuart, K. 1996. RNA editing: a mechanism for gRNA-specified uridylate insertion into precursor mRNA. *Science* **273**: 1189-1195.

- Kang, X., Rogers, K., Gao, G., Falick, A.M., Zhou, S., and Simpson, L. 2005. Reconstitution of uridine-deletion precleaved RNA editing with two recombinant enzymes. *Proc Natl Acad Sci U S A* **102**: 1017-1022.
- Kapushoc, S.T., and Simpson, L. 1999. In vitro uridine insertion RNA editing mediated by cis-acting guide RNAs. *RNA* **5**: 656-669.
- Kiss, T. 2001. Small nucleolar RNA-guided post-transcriptional modification of cellular RNAs. *Embo J* **20**: 3617-3622.
- Kolb, F.A., Westhof, E., Ehresmann, B., Ehresmann, C., Wagner, E.G., and Romby, P. 2001. Four-way junctions in antisense RNA-mRNA complexes involved in plasmid replication control: a common theme? *J Mol Biol* **309**: 605-614.
- Kolb, F.A., Westhof, E., Ehresmann, C., Ehresmann, B., Wagner, E.G., and Romby, P. 2001. Bulged residues promote the progression of a loop-loop interaction to a stable and inhibitory antisense-target RNA complex. *Nucleic Acids Res* **29**: 3145-3153.
- Koller, J., Muller, U.F., Schmid, B., Missel, A., Kruff, V., Stuart, K., and Goring, H.U. 1997. Trypanosoma brucei gBP21. An arginine-rich mitochondrial protein that binds to guide RNA with high affinity. *J Biol Chem* **272**: 3749-3757.
- Koller, J., Norskau, G., Paul, A.S., Stuart, K., and Goring, H.U. 1994. Different Trypanosoma brucei guide RNA molecules associate with an identical complement of mitochondrial proteins in vitro. *Nucleic Acids Res* **22**: 1988-1995.
- Koslowsky, D., Reifur, L, Yu, L, and Chen, W. 2004. Evidence for U-tail Stabilization of gRNA/mRNA Interactions in Kinetoplastid RNA Editing. *RNA biology* **1**: 28-34.
- Koslowsky, D.J. 2004. A historical perspective on RNA editing: how the peculiar and bizarre became mainstream. *Methods Mol Biol* **265**: 161-197.
- Koslowsky, D.J., Bhat, G.J., Perrollaz, A.L., Feagin, J.E., and Stuart, K. 1990. The MURF3 gene of T. brucei contains multiple domains of extensive editing and is homologous to a subunit of NADH dehydrogenase. *Cell* **62**: 901-911.
- Koslowsky, D.J., Bhat, G.J., Read, L.K., and Stuart, K. 1991. Cycles of progressive realignment of gRNA with mRNA in RNA editing. *Cell* **67**: 537-546.
- Koslowsky, D.J., Riley, G.R., Feagin, J.E., and Stuart, K. 1992. Guide RNAs for transcripts with developmentally regulated RNA editing are present in both life cycle stages of Trypanosoma brucei. *Mol Cell Biol* **12**: 2043-2049.
- Koslowsky, D.J., and Yahampath, G. 1997. Mitochondrial mRNA 3' cleavage/polyadenylation and RNA editing in Trypanosoma brucei are independent events. *Mol Biochem Parasitol* **90**: 81-94.

- Krek, A., Grun, D., Poy, M.N., Wolf, R., Rosenberg, L., Epstein, E.J., MacMenamin, P., da Piedade, I., Gunsalus, K.C., Stoffel, M., et al. 2005. Combinatorial microRNA target predictions. *Nat Genet* **37**: 495-500.
- Lai, D.H., Hashimi, H., Lun, Z.R., Ayala, F.J., and Lukes, J. 2008. Adaptations of *Trypanosoma brucei* to gradual loss of kinetoplast DNA: *Trypanosoma equiperdum* and *Trypanosoma evansi* are petite mutants of *T. brucei*. *Proc Natl Acad Sci USA*.
- Lambert, L., Muller, U.F., Souza, A.E., and Goring, H.U. 1999. The involvement of gRNA-binding protein gBP21 in RNA editing-an in vitro and in vivo analysis. *Nucleic Acids Res* **27**: 1429-1436.
- Law, J.A., O'Hearn, S., and Sollner-Webb, B. 2007. In *Trypanosoma brucei* RNA editing, TbMP18 (band VII) is critical for editosome integrity and for both insertional and deletional cleavages. *Mol Cell Biol* **27**: 777-787.
- Law, J.A., O'Hearn, S.F., and Sollner-Webb, B. 2008. *Trypanosoma brucei* RNA editing protein TbMP42 (band VI) is crucial for the endonucleolytic cleavages but not the subsequent steps of U-deletion and U-insertion. *Rna*.
- Lawson, S.D., Igo, R.P., Jr., Salavati, R., and Stuart, K.D. 2001. The specificity of nucleotide removal during RNA editing in *Trypanosoma brucei*. *RNA* **7**: 1793-1802.
- Leegwater, P., Speijer, D., and Benne, R. 1995. Identification by UV cross-linking of oligo(U)-binding proteins in mitochondria of the insect trypanosomatid *Crithidia fasciculata*. *Eur J Biochem* **227**: 780-786.
- Leung, S.S., and Koslowsky, D.J. 1999. Mapping contacts between gRNA and mRNA in trypanosome RNA editing. *Nucleic Acids Res* **27**: 778-787.
- Leung, S.S., and Koslowsky, D.J. 2001a. Interactions of mRNAs and gRNAs involved in trypanosome mitochondrial RNA editing: structure probing of an mRNA bound to its cognate gRNA. *RNA* **7**: 1803-1816.
- Leung, S.S., and Koslowsky, D.J. 2001b. RNA editing in *Trypanosoma brucei*: characterization of gRNA U-tail interactions with partially edited mRNA substrates. *Nucleic Acids Res* **29**: 703-709.
- Levine, E., Zhang, Z., Kuhlman, T., and Hwa, T. 2007. Quantitative characteristics of gene regulation by small RNA. *PLoS Biol* **5**: e229.
- Lewis, B.P., Burge, C.B., and Bartel, D.P. 2005. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **120**: 15-20.
- Lukes, J., Hashimi, H., and Zikova, A. 2005. Unexplained complexity of the mitochondrial genome and transcriptome in kinetoplastid flagellates. *Curr Genet* **48**: 277-299.

- Lunde, B.M., Moore, C., and Varani, G. 2007. RNA-binding proteins: modular design for efficient function. *Nat Rev Mol Cell Biol* **8**: 479-490.
- MacRae, I.J., Ma, E., Zhou, M., Robinson, C.V., and Doudna, J.A. 2008. In vitro reconstitution of the human RISC-loading complex. *Proc Natl Acad Sci USA* **105**: 512-517.
- Madison-Antenucci, S., Grams, J., and Hajduk, S.L. 2002. Editing machines: the complexities of trypanosome RNA editing. *Cell* **108**: 435-438.
- Madison-Antenucci, S., and Hajduk, S.L. 2001. RNA editing-associated protein 1 is an RNA binding protein with specificity for predated mRNA. *Mol Cell* **7**: 879-886.
- Maga, J.A., and LeBowitz, J.H. 1999. Unravelling the kinetoplastid paraflagellar rod. *Trends Cell Biol* **9**: 409-413.
- Maslov, D.A., and Simpson, L. 1992. The polarity of editing within a multiple gRNA-mediated domain is due to formation of anchors for upstream gRNAs by downstream editing. *Cell* **70**: 459-467.
- Mathers, C.D., Ezzati, M., and Lopez, A.D. 2007. Measuring the burden of neglected tropical diseases: the global burden of disease framework. *PLoS Negl Trop Dis* **1**: e114.
- Matthews, K.R. 2005. The developmental cell biology of *Trypanosoma brucei*. *J Cell Sci* **118**: 283-290.
- Mattick, J.S., and Makunin, I.V. 2006. Non-coding RNA. *Hum Mol Genet* **15 Spec No 1**: R17-29.
- McGrail, J.C., and O'Keefe, R.T. 2007. The U1, U2 and U5 snRNAs crosslink to the 5' exon during yeast pre-mRNA splicing. *Nucleic Acids Res.*
- McManus, M.T., Shimamura, M., Grams, J., and Hajduk, S.L. 2001. Identification of candidate mitochondrial RNA editing ligases from *Trypanosoma brucei*. *RNA* **7**: 167-175.
- Michels, P.A., Bringaud, F., Herman, M., and Hannaert, V. 2006. Metabolic functions of glycosomes in trypanosomatids. *Biochim Biophys Acta* **1763**: 1463-1477.
- Militello, K.T., and Read, L.K. 1999. Coordination of kRNA editing and polyadenylation in *Trypanosoma brucei* mitochondria: complete editing is not required for long poly(A) tract addition. *Nucleic Acids Res* **27**: 1377-1385.
- Militello, K.T., Refour, P., Comeaux, C.A., and Duraisingh, M.T. 2008. Antisense RNA and RNAi in protozoan parasites: Working hard or hardly working? *Mol Biochem Parasitol* **157**: 117-126.

- Miller, M.M., Halbig, K., Cruz-Reyes, J., and Read, L.K. 2006. RBP16 stimulates trypanosome RNA editing in vitro at an early step in the editing reaction. *RNA* **12**: 1292-1303.
- Missel, A., Lambert, L., Norskau, G., and Goring, H.U. 1999. DEAD-box protein HEL64 from *Trypanosoma brucei*: subcellular localization and gene knockout analysis. *Parasitol Res* **85**: 324-330.
- Missel, A., Souza, A.E., Norskau, G., and Goring, H.U. 1997. Disruption of a gene encoding a novel mitochondrial DEAD-box protein in *Trypanosoma brucei* affects edited mRNAs. *Mol Cell Biol* **17**: 4895-4903.
- Muller, U.F., and Goring, H.U. 2002. Mechanism of the gBP21-mediated RNA/RNA annealing reaction: matchmaking and charge reduction. *Nucleic Acids Res* **30**: 447-455.
- Muller, U.F., Lambert, L., and Goring, H.U. 2001. Annealing of RNA editing substrates facilitated by guide RNA-binding protein gBP21. *Embo J* **20**: 1394-1404.
- Mulligan, R.M., Chang, K.L., and Chou, C.C. 2007. Computational analysis of RNA editing sites in plant mitochondrial genomes reveals similar information content and a sporadic distribution of editing sites. *Mol Biol Evol* **24**: 1971-1981.
- Naessens, J. 2006. Bovine trypanotolerance: A natural ability to prevent severe anaemia and haemophagocytic syndrome? *Int J Parasitol* **36**: 521-528.
- Nicholson, A.W. 1999. Function, mechanism and regulation of bacterial ribonucleases. *FEMS Microbiol Rev* **23**: 371-390.
- Njiokou, F., Simo, G., Nkinin, S.W., Laveissiere, C., and Herder, S. 2004. Infection rate of *Trypanosoma brucei* s.l., *T. vivax*, *T. congolense* "forest type", and *T. simiae* in small wild vertebrates in south Cameroon. *Acta Trop* **92**: 139-146.
- Nok, A.J. 2003. Arsenicals (melarsoprol), pentamidine and suramin in the treatment of human African trypanosomiasis. *Parasitol Res* **90**: 71-79.
- O'Gorman, G.M., Park, S.D., Hill, E.W., Meade, K.G., Mitchell, L.C., Agaba, M., Gibson, J.P., Hanotte, O., Naessens, J., Kemp, S.J., et al. 2006. Cytokine mRNA profiling of peripheral blood mononuclear cells from trypanotolerant and trypanosusceptible cattle infected with *Trypanosoma congolense*. *Physiol Genomics* **28**: 53-61.
- Ochsenreiter, T., Cipriano, M., and Hajduk, S.L. 2008. Alternative mRNA editing in trypanosomes is extensive and may contribute to mitochondrial protein diversity. *PLoS ONE* **3**: e1566.

- Ochsenreiter, T., and Hajduk, S.L. 2006. Alternative editing of cytochrome c oxidase III mRNA in trypanosome mitochondria generates protein diversity. *EMBO Rep* **7**: 1128-1133.
- Oli, M.W., Cotlin, L.F., Shiflett, A.M., and Hajduk, S.L. 2006. Serum resistance-associated protein blocks lysosomal targeting of trypanosome lytic factor in *Trypanosoma brucei*. *Eukaryot Cell* **5**: 132-139.
- Oppegard, L.M., Hillestad, M., McCarthy, R.T., Pai, R.D., and Connell, G.J. 2003. Cis-acting elements stimulating kinetoplastid guide RNA-directed editing. *J Biol Chem* **278**: 51167-51175.
- Opperdoes, F.R., and Michels, P.A. 2007. Horizontal gene transfer in trypanosomatids. *Trends Parasitol* **23**: 470-476.
- Ou, Y.C., Giroud, C., and Baltz, T. 1991. Kinetoplast DNA analysis of four *Trypanosoma evansi* strains. *Mol Biochem Parasitol* **46**: 97-102.
- Pai, R.D., Oppegard, L.M., and Connell, G.J. 2003. Sequence and structural requirements for optimal guide RNA-directed insertional editing within *Leishmania tarentolae*. *RNA* **9**: 469-483.
- Palazzo, S.S., Panigrahi, A.K., Igo, R.P., Salavati, R., and Stuart, K. 2003. Kinetoplastid RNA editing ligases: complex association, characterization, and substrate requirements. *Mol Biochem Parasitol* **127**: 161-167.
- Panigrahi, A.K., Allen, T.E., Stuart, K., Haynes, P.A., and Gygi, S.P. 2003. Mass spectrometric analysis of the editosome and other multiprotein complexes in *Trypanosoma brucei*. *J Am Soc Mass Spectrom* **14**: 728-735.
- Panigrahi, A.K., Ernst, N.L., Domingo, G.J., Fleck, M., Salavati, R., and Stuart, K.D. 2006. Compositionally and functionally distinct editosomes in *Trypanosoma brucei*. *RNA* **12**: 1038-1049.
- Panigrahi, A.K., Schnauffer, A., Ernst, N.L., Wang, B., Carmean, N., Salavati, R., and Stuart, K. 2003. Identification of novel components of *Trypanosoma brucei* editosomes. *RNA* **9**: 484-492.
- Panigrahi, A.K., Schnauffer, A., and Stuart, K.D. 2007. Isolation and compositional analysis of trypanosomatid editosomes. *Methods Enzymol* **424**: 3-24.
- Parsons, M. 2004. Glycosomes: parasites and the divergence of peroxisomal purpose. *Mol Microbiol* **53**: 717-724.
- Paulus, M., Haslbeck, M., and Watzele, M. 2004. RNA stem-loop enhanced expression of previously non-expressible genes. *Nucleic Acids Res* **32**: e78.

- Pelletier, M., and Read, L.K. 2003. RBP16 is a multifunctional gene regulatory protein involved in editing and stabilization of specific mitochondrial mRNAs in *Trypanosoma brucei*. *RNA* **9**: 457-468.
- Peris, M., Frech, G.C., Simpson, A.M., Bringaud, F., Byrne, E., Bakker, A., and Simpson, L. 1994. Characterization of two classes of ribonucleoprotein complexes possibly involved in RNA editing from *Leishmania tarentolae* mitochondria. *Embo J* **13**: 1664-1672.
- Perron, M.P., and Provost, P. 2008. Protein interactions and complexes in human microRNA biogenesis and function. *Front Biosci* **13**: 2537-2547.
- Piller, K.J., Decker, C.J., Rusche, L.N., Harris, M.E., Hajduk, S.L., and Sollner-Webb, B. 1995. Editing domains of *Trypanosoma brucei* mitochondrial RNAs identified by secondary structure. *Mol Cell Biol* **15**: 2916-2924.
- Pollard, V.W., Harris, M.E., and Hajduk, S.L. 1992. Native mRNA editing complexes from *Trypanosoma brucei* mitochondria. *Embo J* **11**: 4429-4438.
- Pollard, V.W., Rohrer, S.P., Michelotti, E.F., Hancock, K., and Hajduk, S.L. 1990. Organization of minicircle genes for guide RNAs in *Trypanosoma brucei*. *Cell* **63**: 783-790.
- Preall, J.B., and Sontheimer, E.J. 2005. RNAi: RISC gets loaded. *Cell* **123**: 543-545.
- Read, L.K., Goring, H.U., and Stuart, K. 1994. Assembly of mitochondrial ribonucleoprotein complexes involves specific guide RNA (gRNA)-binding proteins and gRNA domains but does not require preedited mRNA. *Mol Cell Biol* **14**: 2629-2639.
- Richardson, N., Navaratnam, N., and Scott, J. 1998. Secondary structure for the apolipoprotein B mRNA editing site. Au-binding proteins interact with a stem loop. *J Biol Chem* **273**: 31707-31717.
- Riley, G.R., Corell, R.A., and Stuart, K. 1994. Multiple guide RNAs for identical editing of *Trypanosoma brucei* apocytochrome b mRNA have an unusual minicircle location and are developmentally regulated. *J Biol Chem* **269**: 6101-6108.
- Riley, G.R., Myler, P.J., and Stuart, K. 1995. Quantitation of RNA editing substrates, products and potential intermediates: implications for developmental regulation. *Nucleic Acids Res* **23**: 708-712.
- Ritchie, W., Legendre, M., and Gautheret, D. 2007. RNA stem-loops: to be or not to be cleaved by RNase III. *RNA* **13**: 457-462.
- Robertson, H.D. 1982. *Escherichia coli* ribonuclease III cleavage sites. *Cell* **30**: 669-672.

- Rosypal, A.C. 2005. Characterization of Canine Leishmaniasis in the United States: Pathogenesis, Immunological Responses, and Transmission of an American Isolate of *Leishmania infantum*. In *Department of Biomedical Sciences and Pathobiology*, pp. 169. Virginia Polytechnic Institute & State University, Virginia.
- Rusche, L.N., Cruz-Reyes, J., Piller, K.J., and Sollner-Webb, B. 1997. Purification of a functional enzymatic editing complex from *Trypanosoma brucei* mitochondria. *Embo J* **16**: 4069-4081.
- Rusche, L.N., Huang, C.E., Piller, K.J., Hemann, M., Wirtz, E., and Sollner-Webb, B. 2001. The two RNA ligases of the *Trypanosoma brucei* RNA editing complex: cloning the essential band IV gene and identifying the band V gene. *Mol Cell Biol* **21**: 979-989.
- Sacharidou, A., Cifuentes-Rojas, C., Halbig, K., Hernandez, A., Dangott, L.J., De Nova-Ocampo, M., and Cruz-Reyes, J. 2006. RNA editing complex interactions with a site for full-round U deletion in *Trypanosoma brucei*. *RNA* **12**: 1219-1228.
- Salavati, R., Ernst, N.L., O'Rear, J., Gilliam, T., Tarun, S., Jr., and Stuart, K. 2006. KREPA4, an RNA binding protein essential for editosome integrity and survival of *Trypanosoma brucei*. *RNA* **12**: 819-831.
- Schmid, B., Read, L.K., Stuart, K., and Goring, H.U. 1996. Experimental verification of the secondary structures of guide RNA-pre-mRNA chimaeric molecules in *Trypanosoma brucei*. *Eur J Biochem* **240**: 721-731.
- Schmid, B., Riley, G.R., Stuart, K., and Goring, H.U. 1995. The secondary structure of guide RNA molecules from *Trypanosoma brucei*. *Nucleic Acids Res* **23**: 3093-3102.
- Schnauffer, A., Clark-Walker, G.D., Steinberg, A.G., and Stuart, K. 2005. The F1-ATP synthase complex in bloodstream stage trypanosomes has an unusual and essential function. *Embo J* **24**: 4029-4040.
- Schnauffer, A., Domingo, G.J., and Stuart, K. 2002. Natural and induced dyskinetoplastic trypanosomatids: how to live without mitochondrial DNA. *Int J Parasitol* **32**: 1071-1084.
- Schnauffer, A., Ernst, N.L., Palazzo, S.S., O'Rear, J., Salavati, R., and Stuart, K. 2003. Separate insertion and deletion subcomplexes of the *Trypanosoma brucei* RNA editing complex. *Mol Cell* **12**: 307-319.
- Schnauffer, A., Panigrahi, A.K., Panicucci, B., Igo, R.P., Jr., Wirtz, E., Salavati, R., and Stuart, K. 2001. An RNA ligase essential for RNA editing and survival of the bloodstream form of *Trypanosoma brucei*. *Science* **291**: 2159-2162.
- Schneider, A. 2001. Unique aspects of mitochondrial biogenesis in trypanosomatids. *Int J Parasitol* **31**: 1403-1415.

- Schumacher, M.A., Karamooz, E., Zikova, A., Trantirek, L., and Lukes, J. 2006. Crystal structures of *T. brucei* MRP1/MRP2 guide-RNA binding complex reveal RNA matchmaking mechanism. *Cell* **126**: 701-711.
- Seed, J.R., and Wenck, M.A. 2003. Role of the long slender to short stumpy transition in the life cycle of the african trypanosomes. *Kinetoplastid Biol Dis* **2**: 3.
- Seiwert, S.D., Heidmann, S., and Stuart, K. 1996. Direct visualization of uridylyate deletion in vitro suggests a mechanism for kinetoplastid RNA editing. *Cell* **84**: 831-841.
- Seiwert, S.D., and Stuart, K. 1994. RNA editing: transfer of genetic information from gRNA to precursor mRNA in vitro. *Science* **266**: 114-117.
- Shikanai, T. 2006. RNA editing in plant organelles: machinery, physiological function and evolution. *Cell Mol Life Sci* **63**: 698-708.
- Shu, H.H., Stuart, K., and Goring, H.U. 1995. Guide RNA molecules not engaged in RNA editing form ribonucleoprotein complexes free of mRNA. *Biochim Biophys Acta* **1261**: 349-359.
- Simpson, A.G., Stevens, J.R., and Lukes, J. 2006. The evolution and diversity of kinetoplastid flagellates. *Trends Parasitol* **22**: 168-174.
- Simpson, A.M., Bakalara, N., and Simpson, L. 1992. A ribonuclease activity is activated by heparin or by digestion with proteinase K in mitochondrial extracts of *Leishmania tarentolae*. *J Biol Chem* **267**: 6782-6788.
- Simpson, L. 1987. The mitochondrial genome of kinetoplastid protozoa: genomic organization, transcription, replication, and evolution. *Annu Rev Microbiol* **41**: 363-382.
- Simpson, L. 1997. The genomic organization of guide RNA genes in kinetoplastid protozoa: several conundrums and their solutions. *Mol Biochem Parasitol* **86**: 133-141.
- Simpson, L. 2001. RNA Editing. *Origin and Evolution of Life* **1**: 501-516.
- Simpson, L., Sbicego, S., and Aphasizhev, R. 2003. Uridine insertion/deletion RNA editing in trypanosome mitochondria: a complex business. *RNA* **9**: 265-276.
- Simpson, L., and Thiemann, O.H. 1995. Sense from nonsense: RNA editing in mitochondria of kinetoplastid protozoa and slime molds. *Cell* **81**: 837-840.
- Simpson, L., Wang, S.H., Thiemann, O.H., Alfonzo, J.D., Maslov, D.A., and Avila, H.A. 1998. U-insertion/deletion Edited Sequence Database. *Nucleic Acids Res* **26**: 170-176.

- Sommer, B., Kohler, M., Sprengel, R., and Seeburg, P.H. 1991. RNA editing in brain controls a determinant of ion flow in glutamate-gated channels. *Cell* **67**: 11-19.
- Souza, A.E., Myler, P.J., and Stuart, K. 1992. Maxicircle CR1 transcripts of *Trypanosoma brucei* are edited and developmentally regulated and encode a putative iron-sulfur protein homologous to an NADH dehydrogenase subunit. *Mol Cell Biol* **12**: 2100-2107.
- Souza, A.E., Shu, H.H., Read, L.K., Myler, P.J., and Stuart, K.D. 1993. Extensive editing of CR2 maxicircle transcripts of *Trypanosoma brucei* predicts a protein with homology to a subunit of NADH dehydrogenase. *Mol Cell Biol* **13**: 6832-6840.
- Sternberg, J.M. 2004. Human African trypanosomiasis: clinical presentation and immune response. *Parasite Immunol* **26**: 469-476.
- Stoltzfus, A. 1999. On the possibility of constructive neutral evolution. *J Mol Evol* **49**: 169-181.
- Stuart, K., and Panigrahi, A.K. 2002. RNA editing: complexity and complications. *Mol Microbiol* **45**: 591-596.
- Stuart, K., Panigrahi, A.K., Schnauffer, A., Drozdz, M., Clayton, C., and Salavati, R. 2002. Composition of the editing complex of *Trypanosoma brucei*. *Philos Trans R Soc Lond B Biol Sci* **357**: 71-79.
- Stuart, K., Salavati, R., Igo, R.P., Lewis Ernst, N., Palazzo, S.S., and Wang, B. 2004. In vitro assays for kinetoplastid U insertion-deletion editing and associated activities. *Methods Mol Biol* **265**: 251-272.
- Stuart, K.D., Schnauffer, A., Ernst, N.L., and Panigrahi, A.K. 2005. Complex management: RNA editing in trypanosomes. *Trends Biochem Sci* **30**: 97-105.
- Sturm, N.R., Maslov, D.A., Blum, B., and Simpson, L. 1992. Generation of unexpected editing patterns in *Leishmania tarentolae* mitochondrial mRNAs: misediting produced by misguiding. *Cell* **70**: 469-476.
- Sturm, N.R., and Simpson, L. 1990. Partially edited mRNAs for cytochrome b and subunit III of cytochrome oxidase from *Leishmania tarentolae* mitochondria: RNA editing intermediates. *Cell* **61**: 871-878.
- Tarun, S.Z., Jr., Schnauffer, A., Ernst, N.L., Proff, R., Deng, J., Hol, W., and Stuart, K. 2007. KREPA6 is an RNA-binding protein essential for editosome integrity and survival of *Trypanosoma brucei*. *RNA*.
- Tomizawa, J. 1986. Control of ColE1 plasmid replication: binding of RNA I to RNA II and inhibition of primer formation. *Cell* **47**: 89-97.

- Trotter, J.R., Ernst, N.L., Carnes, J., Panicucci, B., and Stuart, K. 2005. A deletion site editing endonuclease in *Trypanosoma brucei*. *Mol Cell* **20**: 403-412.
- van Weelden, S.W., van Hellemond, J.J., Opperdoes, F.R., and Tielens, A.G. 2005. New functions for parts of the krebs cycle in procyclic *trypanosoma brucei*, a cycle not operating as a cycle. *J Biol Chem*.
- Vanhamme, L., Perez-Morga, D., Marchal, C., Speijer, D., Lambert, L., Geuskens, M., Alexandre, S., Ismaili, N., Goringe, U., Benne, R., et al. 1998. *Trypanosoma brucei* TBRGG1, a mitochondrial oligo(U)-binding protein that co-localizes with an in vitro RNA editing activity. *J Biol Chem* **273**: 21825-21833.
- Vondruskova, E., van den Burg, J., Zikova, A., Ernst, N.L., Stuart, K., Benne, R., and Lukes, J. 2004. RNA interference analyses suggest a transcript-specific regulatory role for MRP1 and MRP2 in RNA editing and other RNA processing in *trypanosoma brucei*. *J Biol Chem*.
- Vondruskova, E., van den Burg, J., Zikova, A., Ernst, N.L., Stuart, K., Benne, R., and Lukes, J. 2005. RNA interference analyses suggest a transcript-specific regulatory role for mitochondrial RNA-binding proteins MRP1 and MRP2 in RNA editing and other RNA processing in *Trypanosoma brucei*. *J Biol Chem* **280**: 2429-2438.
- Wakeman, C.A., Winkler, W.C., and Dann, C.E., 3rd. 2007. Structural features of metabolite-sensing riboswitches. *Trends Biochem Sci* **32**: 415-424.
- Walter, N.G., and Engelke, D.R. 2002. Ribozymes: catalytic RNAs that cut things, make things, and do odd and useful jobs. *Biologist (London)* **49**: 199-203.
- Wang, B., Ernst, N.L., Palazzo, S.S., Panigrahi, A.K., Salavati, R., and Stuart, K. 2003. TbMP44 is essential for RNA editing and structural integrity of the editosome in *Trypanosoma brucei*. *Eukaryot Cell* **2**: 578-587.
- WHO. 2001. Report of the Scientific Working Group meeting on African trypanosomiasis, pp. 182, Geneva.
- Worthey, E.A., Schnauffer, A., Mian, I.S., Stuart, K., and Salavati, R. 2003. Comparative analysis of editosome proteins in trypanosomatids. *Nucleic Acids Res* **31**: 6392-6408.
- Xia, T., SantaLucia, J., Jr., Burkard, M.E., Kierzek, R., Schroeder, S.J., Jiao, X., Cox, C., and Turner, D.H. 1998. Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs. *Biochemistry* **37**: 14719-14735.
- Yu, L.E. 2006. Elucidating the structure and kinetics of the apocytochrome B mRNA/gRNA complex in *Trypanosoma brucei* mitochondria. In *Cell and Molecular Biology*, pp. 205. Michigan State University, East Lansing.

- Yu, L.E., and Koslowsky, D.J. 2006. Interactions of mRNAs and gRNAs involved in trypanosome mitochondrial RNA editing: structure probing of a gRNA bound to its cognate mRNA. *RNA* **12**: 1050-1060.
- Zambrano-Villa, S., Rosales-Borjas, D., Carrero, J.C., and Ortiz-Ortiz, L. 2002. How protozoan parasites evade the immune response. *Trends Parasitol* **18**: 272-278.
- Zeiler, B.N., and Simons, R.W. 1998. *Antisense RNA Structure and Function*. Cold Spring Harbor Laboratory Press, New York, pp. 741.
- Zhang, K., and Nicholson, A.W. 1997. Regulation of ribonuclease III processing by double-helical sequence antideterminants. *Proc Natl Acad Sci USA* **94**: 13437-13441.
- Zikova, A., Horakova, E., Jirku, M., Dunajcikova, P., and Lukes, J. 2006. The effect of down-regulation of mitochondrial RNA-binding proteins MRP1 and MRP2 on respiratory complexes in procyclic *Trypanosoma brucei*. *Mol Biochem Parasitol* **149**: 65-73.

## Chapter 2

### ***TRYPANOSOMA BRUCEI* ATPASE SUBUNIT 6 MRNA BOUND TO GA6-14 FORMS A CONSERVED THREE- HELICAL STRUCTURE**

Reifur, Larissa; Koslowsky, Donna J. (2008). *Trypanosoma brucei* ATPase subunit 6 mRNA bound to gA6-14 forms a conserved three-helical structure RNA 2008 14: 2195-2211. doi:10.1261/ma.1144508

## **ABSTRACT**

*T. brucei* survival relies on the expression of mitochondrial genes, most of which require RNA editing to become translatable. In trypanosomes, RNA editing involves the insertion and deletion of uridylates, a developmentally regulated process directed by guide RNAs (gRNAs) and catalyzed by the editosome, a complex of proteins. The pathway for mRNA/gRNA complex formation and assembly with the editosome is still unknown. Work from our laboratory has suggested that distinct mRNA/gRNA complexes anneal to form a conserved core structure that may be important for editosome assembly. The secondary structure for the apocytochrome b (CYb) pair has been previously determined and corroborates our model of a three-helical structure. Here, we used crosslinking and solution structure probing experiments to determine the structure of the ATPase subunit 6 (A6) mRNA hybridized to its cognate gA6-14 gRNA in different stages of editing. Our results indicate that both unedited and partially edited A6/gA6-14 pairs fold into a three-helical structure similar to the previously characterized CYb/gCYb-558 pair. These results lead us to conclude that at least two mRNA/gRNA pairs with distinct editing sites and distinct primary sequences fold to a three-helical secondary configuration that persists through the first few editing events.

## **INTRODUCTION**

*Trypanosoma brucei* is a protozoan parasite that undergoes mitochondrial RNA editing to survive. This post-transcriptional modification is unique to kinetoplastids and involves site-specific uridylate (U) insertions and deletions to correct encoded frameshifts, create translation start and stop codons, and extend open reading frames (Simpson et al. 2003;

Stuart et al. 2005). In *T. brucei*, 12 of the 18 mRNAs require editing, but the extent and control of the modification vary according to the message and life cycle of the parasite. Short (50-70 nucleotides [nt]) guide RNAs (gRNA) with complementarity to segments of pre-edited or partially edited mRNAs provide the sequence information necessary for the precise modifications (Blum et al. 1990). Each gRNA has three functionally distinct domains that seem to fold onto a similar structure containing one to two stem-loops (Schmid et al. 1995; Yu and Koslowsky 2006). The gRNA anchor domain at the 5'-end (4-16 nt) is complementary to the anchor binding site (ABS) within the mRNA. Hybridization of gRNA anchor and ABS forms the anchor helix, which is fundamental for editing initiation. The first mismatch 5' of the anchor helix (mRNA orientation) is the presumed first editing site (ES). At the opposing gRNA end, there is the 3'-oligo (U) tail (U-tail) domain (5-24 nt) that is added post-transcriptionally and thought to tether the purine rich region of the mRNA and form the U-tail helix (Blum and Simpson 1990; Seiwert et al. 1996; Leung and Koslowsky 1999; 2001b). Joining the anchor helix and the U-tail helix is the central portion of the gRNA, or guiding domain, which dictates the type of editing (insertion or deletion) and the number of U's involved in the process. Annealing of the gRNA to the mRNA is proposed to form a three-helical structure that defines the editing site (Leung and Koslowsky 2001a; 2001b; Yu and Koslowsky 2006) and allows editing to proceed mostly from 3' to 5' within the mRNA (Abraham et al. 1988; Sturm and Simpson 1990; Koslowsky et al. 1991; Maslov and Simpson 1992).

The three-helical structure was not initially predicted to exist in distinct mRNA/gRNA pairs. Studies using computer programs to obtain secondary structures for different mRNA/gRNA hybrids found no consensus secondary structure among the

complexes (Blum and Simpson 1990; Leung and Koslowsky 1999). Only after crosslinking three gRNAs to their respective cognate mRNAs and incorporating the resulting constraints into a computer modeling program were all three pairs shown to fold into a three-helical structure (Leung and Koslowsky 1999) similar to the structure of a fourth mRNA/gRNA pair confirmed previously by S1 protection assay (Blum and Simpson 1990). Through crosslinking and structure probing studies, a secondary structure model has been proposed for a CYb mRNA/gRNA pair (Leung and Koslowsky 2001a; 2001b; Yu and Koslowsky 2006), and it shows the persistence of three helices even after editing of the first sites, raising the importance of this organization.

Editing involves a cascade of coordinated steps catalyzed by the editosome, a multi-protein complex that can be isolated and enriched from kinetoplastid mitochondria (Pollard et al. 1992; Corell et al. 1996; Rusche et al. 1997; Simpson et al. 2004). Editosomes can be divided into insertion and deletion subcomplexes (Schnauffer et al. 2003; Panigrahi et al. 2006), which optimally work under distinct conditions (Cruz-Reyes et al. 1998). A third editosome has been found to specifically cleave insertion sites when the gRNA appears in cis (Panigrahi et al. 2007; Carnes et al. 2008). Each editosome has unique proteins such as a specific endonuclease and a structural protein. Nonetheless, all subcomplexes share numerous structural and catalytic proteins and have the three enzymatic activities required for editing: endonuclease, terminal uridylyl transferase or 3' exonuclease, and ligase. In addition, accessory factors, such as RBP16 and the MRP complex, have been shown to transiently interact with the editosome and indirectly affect editing (Koller et al. 1997; Blom et al. 2001; Aphasizhev et al. 2003; Panigrahi et al. 2003; Vondruskova et al. 2005; Miller et al. 2006; Ammerman et al. 2008; Zikova et al.

2008). Although RBP16 and the MRP complex have been shown to facilitate the annealing of a few mRNA/gRNAs, in vitro full-round and precleaved editing assays are conducted with purified editosomes in the absence of these or other proteins (Seiwert and Stuart 1994; Cruz-Reyes and Sollner-Webb 1996; Igo et al. 2000; Kang et al. 2005).

Despite all the progress, there are unanswered questions concerning the temporal association of gRNAs and mRNAs with the editosome and what triggers the editing cascade. It has been suggested that gRNAs and mRNAs are initially bound to protein complexes and then brought together to assemble into a binary complex away from the editosome (Zikova et al. 2008) or within the editosome (Madison-Antenucci et al. 2002). Alternatively, the mRNA/gRNA complexes could form in the absence of proteins and be the signal for editosome recognition and editing (Blum and Simpson 1990; Rusche et al. 1997; Leung and Koslowsky 1999; Yu and Koslowsky 2006).

The three-helical model proposed by our laboratory is believed to fit several mRNA/gRNA pairs. Here we determined the secondary structure of the A6/gA6-14 pair during different stages of editing using both cross-linking and solution structure probing methods. Like the previously described CYb pair, the A6 pair folds into a three-helical structure that is maintained during the initial stages of editing. However, in contrast to the CYb pair that only undergoes insertional editing, the A6 pairs studied had only editing sites of the deletion type. The highly organized structure observed within the editing sites in A6 and CYb binary complexes may provide new insights into a mechanism for editing site specificity.

## **RESULTS**

### **Description of RNAs**

The ATPase subunit 6 (A6) transcript is extensively and constitutively edited throughout the parasite's life cycle (Bhat et al. 1990). The mRNA used in this study contains the natural sequence corresponding to one domain from the 3'-end of *T. brucei* A6, which is a normal target of editing. We used both the unedited A6U and the partially edited through 3 sites A6P3 substrates (Fig. 2.1). Editing of this domain is initiated by gA6-14 and both the mRNA and gRNA have been previously described (Leung and Koslowsky 1999) and extensively used in editing assays in vitro (Seiwert and Stuart 1994; Cruz-Reyes and Sollner-Webb 1996; Kable et al. 1996; Seiwert et al. 1996). While the 5'-end of the gRNA specifically binds the mRNA ABS, and its middle portion directs editing, the role of the gRNAs 3'-end, or U-tail, is less clear. It is hypothesized to tether the purine-rich region upstream of the mRNA editing site (Blum and Simpson 1990; Leung and Koslowsky 1999) facilitating ligation of the mRNA pieces after editing (Seiwert et al. 1996). Mutations strengthening U-tail binding improved U insertion (Burgess et al. 1999; Kapushoc and Simpson 1999; Igo et al. 2000) whereas U deletions were inhibited in some experiments (Seiwert et al. 1996) but increased in others (Cruz-Reyes et al. 2001). This can be explained by the fact that the U-tail differentially contributes to the binding affinity of distinct mRNA/gRNA pairs (Koslowsky et al. 2004). Because of the curious influence of the U-tail on editing, we included in our experiments two variations of gA6-14; one with deleted U-tail (gA6-14sU) and another with 6 U-to-C mutations (gA6-14-C6) that strengthens U-tail binding. gA6-14-C6 has been previously published as g[2,1]+6C and it was shown to increase deletion editing efficiency fourfold when compared to gA6-14 (Cruz-Reyes et al. 2001).

**A6U**

5' g<sub>-51</sub>gcgaattgggtaccGGAAUUGCCUUUGCCAAACUUUUAGAAGAAAGAGC  
 A<sub>1</sub>GGAAAGGUUAGGGGGAGGAGAGAAGAAAGGGAAAGUUGAUUU<sup>**U<sub>45</sub>UGGAGUUA**</sup>  
<sup>**UAG<sub>56</sub>**</sup>AAUAAGAUCAAAUAAGUUAUAUAUA<sub>81</sub>3'

**A6P3**

5' g<sub>-51</sub>ggcgaattgggtaccGGAAUUGCCUUUGCCAAACUUUUAGAAGAAAGAGC  
 A<sub>1</sub>GGAAAGGUUAGGGGGAGGAGAGAAGAAAGGGAAAGUUG<sup>**G<sub>39</sub>UuuGuuA\* \*UUGGA**</sup>  
<sup>**GUUAUAG<sub>58</sub>**</sup>AAUAAGAUCAAAUAAGUUAUAUAUA<sub>83</sub>3'

**gA6-14**

5' G<sub>1</sub>GACUAUAACUCCGAUAACGAAUCAGAUUUUGACAGUGAUUAUGAUAAUUA<sub>50</sub>U  
 UUUUUUUUUUUUUU<sub>65</sub>3'

**sU**: without U-tail

**C6**: 5' UUUCCCUUUCUUCUC3'

Figure 2.1 A6 mRNAs and gRNAs.

Both unedited (A6U) and partially edited through 3 sites (A6P3) mRNAs are highlighted at the anchor binding site (ABS), which is complementary to the 5'-end (anchor) of gA6-14. A6P3 differs from A6U in the deletion of two uridylates at ES1 (asterisks) and 4 inserted uridylates at ES2 and ES3 (highlighted lower case). In order to coordinate the numbering of the different substrates used, the shorter mRNA sequences were numbered starting with the 5'-most nucleotide as number 1 or A<sub>1</sub>. Nucleotides found upstream of A<sub>1</sub> in the longer substrates, were negative numbers (up to G<sub>51</sub>). The underlined sequence within gA6-14 indicates the U-tail. The suffix sU for gA6-14 indicates absence of U-tail and C6 corresponds to modified U-tail. The mRNA nucleotides in lower case at the 5'-end are vector sequence.

## gRNA U-tail interactions with mRNA

### *Cross-linking of mRNA/gRNA*

In vivo, gRNA U-tails are heterogeneous in size averaging 15 U's (Blum and Simpson

1990). In vitro synthesized gRNA containing a U-tail with 10 U's (U<sub>10</sub>-tail) had

previously been used in cross-link studies to map the U-tail binding site within different

mRNAs. Leung and Koslowsky (1999; 2001b) attached the azidophenacyl bromide

(APA) cross-linking agent to the 3'-end of the gRNA containing a U<sub>10</sub>-tail and cross-

linked the gRNA to the mRNA with UV irradiation. They observed that the U-tail

interacted with the purine-rich region (4-15 nt) upstream of the editing site using three distinct mRNA/gRNA pairs. With the purpose of having an mRNA/gRNA interaction closer to what is seen in vivo, in this study we used a U<sub>15</sub>-tail and the mRNA with native sequence. The U-tail was chemically synthesized (Dharmacon Research) containing the cross-linking agent 4-thio-U group (S<sup>4</sup>U) at the 5th or 10th U, counting from the 5'-end. These U5 and U10 modified tails were 5'-end radioactively labeled with [ $\gamma$ -<sup>32</sup>P] ATP, ligated to gA6-14sU (hence called gA6-14-U5 and gA6-14-U10), and gel purified. Half of the recovered samples received further treatment with azidophenacyl bromide (APA). APA is coupled to the thio group to increase the range of cross-links from 0 Å to 9 Å (Thomas et al. 2000). The gRNAs, gA6-14-U5 and gA6-14-U10, with S<sup>4</sup>U or APA, were then hybridized to A6U and A6P3 under editing conditions, followed by UV irradiation for 20 min on ice as previously described (Burgin and Pace 1990; Leung and Koslowsky 1999). Through denaturing gel purification, we observed that the cross-linked (x) molecules for both mRNAs divided into two populations (B1 and B2) with distinct mobilities (Fig. 2.2). B1 cross-links were more efficient than B2, except for A6U x gA6-14-U10 (see explanation under Primer extension of cross-linked mRNA/gRNA, below). Regarding the abundance of cross-linked molecules, addition of the APA group increased the cross-linking efficiency in comparison to S<sup>4</sup>U, as expected. S<sup>4</sup>U is a zero length agent that only efficiently cross-links to single stranded regions, a characteristic that decreases the cross-linking efficiency but generates precise results (Dubreuil et al. 1991; Favre et al. 1998). While the addition of APA increased the amount of B1 over fivefold, only a modest increase was observed in the B2 population.

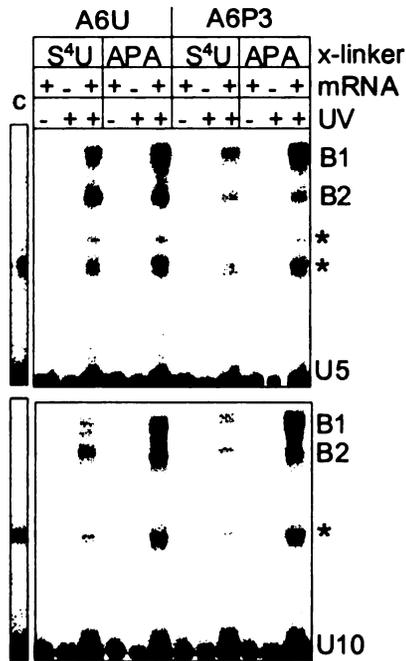


Figure 2.2 A6U and A6P3 crosslinks with 5' <sup>32</sup>P-labeled gA6-14-U5, (U5) or gA6-14-U10, (U10). U5 and U10 indicate that the position of the cross-linking agent (S<sup>4</sup>U or APA) was at the 5th or 10th uridylyte, respectively, in the U-tail. B1 and B2 are two populations of cross-linked molecules. Asterisks represent intra-gRNA cross-links that were also present in the control lane (no mRNA) when the gel was over exposed (representative lane, labeled C) or were not consistently observed in all experiments. Due to limited amounts of sample, only 1/10 of the cross-linked molecules were used in the control lanes (no mRNA and no UV lanes).

#### *Primer extension of cross-linked mRNA/gRNA*

The cross-linked molecules were gel purified and analyzed by primer extension as previously described (Leung and Koslowsky 1999). The reverse transcriptase (RT) is halted one base prior to the cross-link, aborting a product that can be mapped on a denaturing polyacrylamide gel. This procedure allows identification of the cross-link site with single nucleotide definition (Sontheimer 1994). As previously reported, we observed bands in the controls not exposed to UV irradiation (data not shown) but they were artifacts due to premature stoppage of the RT at structural elements in the mRNA sequence (Sontheimer 1994). Thus, a true cross-linked site was considered only when the band was stronger than the corresponding band on the control in multiple experiments. In addition to primer extension, we conducted ribonuclease (RNase) H assays to confirm identified cross-links, as recommended by Sontheimer (1994). APA augmented both the cross-linking and RT halting efficiency and facilitated subsequent mapping; however, it cross-linked to a broad number of nucleotides. With S<sup>4</sup>U alone, the

percentage of cross-linked molecules was low and they were more difficult to map; however, the results were most often supportive, overlapping with the APA sites. A summary of the mapped cross-linked sites is shown in Figure 2.3.

Identified cross-link sites in B1, with or without APA, generally located upstream of the first few editing sites. For both A6U and A6P3, the strongest mapped cross-link sites were complementary for both the U5 and U10 thio placements; U5 cross-linked most strongly with nucleotides 29-33 and U10 with 23-28. In the absence of APA, the gRNA with the S<sup>4</sup>U in the U5 position did cross-link to additional sites closer to the ABS. Strong U5-S<sup>4</sup>U cross-links were observed at positions 37 and 39 within the unedited mRNA and weak, but distinct cross-links were observed at positions 38 and 39 within the partially edited substrate. Although both of these sites are predicted to be within a helical region in our model, they were located at the end of a stem and may reflect the propensity for S<sup>4</sup>U to prefer unpaired bases. With the addition of APA, gA6-14-U10 also induced weak but distinct cross-links at nucleotides 7 and 8 in both the unedited and partially edited substrates. This cross-link site was not identified with gA6-14-U5 through primer extension; however, the RNase H analyses (presented below) suggest that similar cross-links did exist when the APA was in U5. Unexpectedly, when the A6P3 substrate was paired with either the U5 or U10 thio-modified gRNA, distinct cross-links were also detected within the anchor helix at nucleotides 51-54. This 51-54 cross-link site was not identified in A6U through primer extension; however, the RNase H analyses (presented below) suggest that similar cross-links did exist in the unedited substrate.

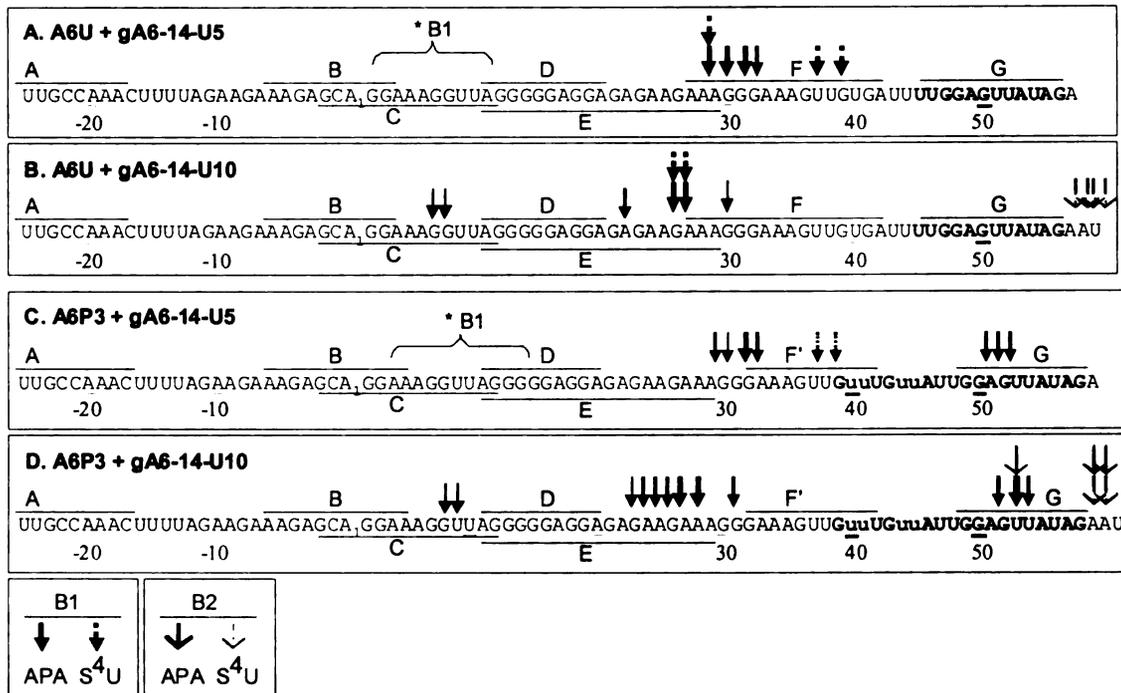


Figure 2.3 Mapped crosslink sites within A6U and A6P3 crosslinked to gA6-14-U5 or gA6-14-U10 (gRNA sequence is not shown). Arrows indicate sites determined through primer extension and confirmed through RNase H analysis. Arrow thickness correlates with intensity of the RT stop. Lines *above* and *below* the sequence labeled A to G indicate the position of each oligodeoxynucleotide (ODN) used in RNase H assays. Sequence numbering is according to Figure 2.1. \*B1 indicates approximate position of cross-link site detected by RNase H assay only. B1 (solid head arrows) and B2 (open head arrows) are cross-link populations shown in Figure 2.2.

U-tail cross-linking sites within the B2 complexes were much more difficult to map by primer extension. Analyses of A6U and A6P3 cross-linked to gA6-14-U5 showed no distinct transcription stops. Analyses of A6U and A6P3 cross-linked to gA6-14-U10 however, indicated cross-linking to nucleotides located within or just downstream of the ABS. This suggests that a tertiary folding may place the U-tail helix closer to the anchor helix.

As previously reported, reverse transcriptase has difficulty transcribing through specific residues within the A6 mRNA sequence (Leung and Koslowsky 1999). The A<sub>3</sub>G<sub>3</sub>A<sub>3</sub> sequence located at nucleotides 27-35 appears to be particularly difficult. Since the majority of our mapped cross-link sites fell within this region, RNase H mapping was used to confirm their presence. Although RNase H mapping is less precise than RT, it can confirm the existence of a cross-link within a small sequence between two oligodeoxynucleotides (ODNs) (Sontheimer 1994). In addition, the sensitivity of the RNase H assay allows the detection of low-population cross-links. Our assays involved hybridizing the radiolabeled cross-linked mRNA/gRNA molecules to a combination of 7 ODNs (Fig. 2.3) and digesting with RNase H. The ODN positions within the mRNAs are shown in Figure 2.3 and representative RNase H digests are in Figure 2.4. Because the radiolabel is carried within the gRNA, only fragments containing a cross-link are visible.

Using the RNase H technique, the presence of strong B1 cross-links between nucleotides 22 and 39 within A6U and A6P3 substrates were confirmed (Fig. 2.4A, B).

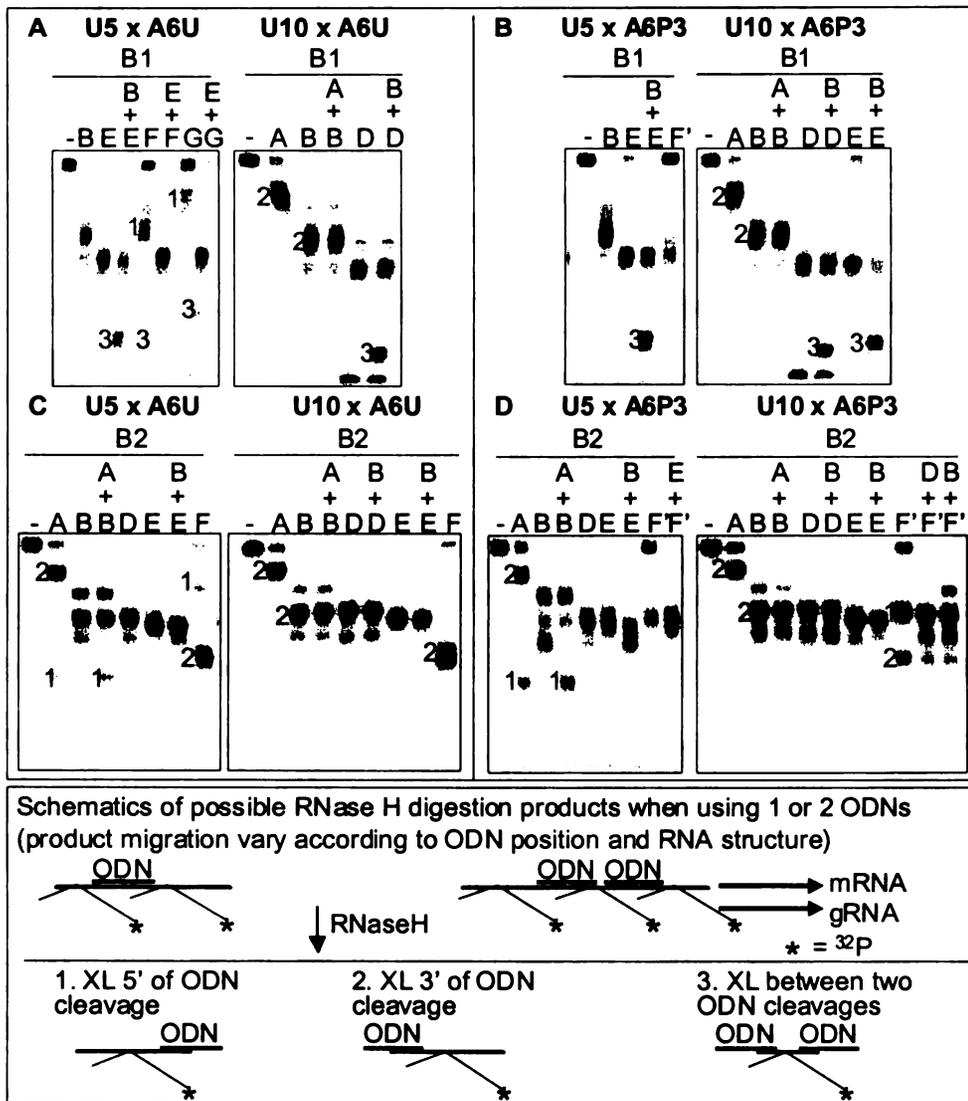


Figure 2.4 RNase H assays confirming crosslink sites for B1 and B2 populations of A6U and A6P3 crosslinked to gA6-14-U5 (U5) or gA6-14-U10 (U10). *Top* panels are B1 cross-links with A6U (A) and A6P3 (B). *Bottom* panels are B2 cross-links with A6U (C) and A6P3 (D). (A-G) ODNs complementary to specific sequence within the mRNA, specified in Figure 3. (Lanes labeled with minus sign) control without ODN. Note incomplete digestion with ODNs F, G, A, and E, due to poor hybridization to mRNA probably because of potential interfering cross-links or anchor helix formation. This generated less digestion product and therefore, bands 3 are often light. Numbers on the *left* side of bands correlate with digestion products shown in the schematic. Light bands at the *bottom* of each gel lane are radioactively labeled gRNA that detached from the mRNA during the experimental procedure.

RNase H digestions using one ODN at a time confirmed the presence of cross-links downstream of E and upstream of F (F') (Fig. 2.4A, B; data not shown). In addition, digestion with ODNs B+E and B+D resulted in the formation of two labeled products; one the same size as with E alone (cross-links found downstream of E) and one small product (cross-links found between B and E or D). This confirms the RT cross-links at nucleotides 7 and 8 for gA6-14-U10 x A6U or A6P3 and suggests that cross-links at the same site, though not detected by RT, were also present with gA6-14-U5 x A6U and A6P3 (Fig. 2.3, 2.4A, B, asterisks). It may be that these sites were not detected using reverse transcriptase in the U5 cross-links due to the purine-rich nature of this region. Incomplete digestion of the mRNA using ODNs F (F') and G make double digests using these ODNs more difficult to interpret because the intensity of the product band does not correlate with quantity of cross-linked molecules. Incomplete digestion is due to impaired hybridization of these ODNs because the corresponding mRNA site coincides either with the binding site for the gRNA anchor or with cross-link sites. Nonetheless, digestions of A6U cross-links with E+F, E+G or D+F generated two detectable products: one population corresponding to the incomplete digestions (not digested by F or G) and a small product (band 3, cross-links located between the two ODNs), confirming once again the cross-links within nucleotides 22-39. Double digestions using the F' and G ODNs for the A6P3 cross-links were much more difficult to interpret. Digestion with F' alone generated a product that migrated either slightly slower or even with the D and E ODN generated products, suggesting that most of the cross-links were upstream of F'. No smaller product was observed, so we could not confirm the RT identified cross-links

downstream from F'. In addition, double digests with ODNs E and either F' or G were not conclusive (data not shown).

RNase H analyses of the gA6-14-U5 B2 cross-links indicated that this population also contained cross-links in at least two distinct areas (Fig. 2.4C, D). Digestion of these cross-links with ODN A gave rise to two products indicating that the cross-links were located both upstream and downstream from A for both A6U and A6P3. Digestion with multiple ODN combinations (A+C, C+E, E+F, E+G, and D+G [data not shown] and A+B, B+E, B+D, B+F', and D+F') excluded the presence of cross-links between ODNs A and F (F'), indicating that one cross-link population was upstream of A and the second population located downstream from the ABS. Digestion of A6U x gA6-14-U5 with G alone (data not shown) and F alone generated two mobility products supporting the finding of two cross-link populations. Digestion with D or E alone probably gave two fragments of similar size that overlapped in Figure 2.4C. Unfortunately, digestion of A6P3 x gA6-14-U5 (B2) with ODN F' was also incomplete. However, double digestion with ODNs B+E and E+F' generated only one fragment, and A+B generated fragments similar to A and B alone, confirming that the cross-links were located upstream of A and downstream of F'.

Analyses of the gA6-14-U10 B2 cross-links (Fig. 2.3, 2.4C, D) were similar to those observed for U5 with one major exception: cross-links upstream of A were not detected (note absence of band #1 when A was used alone on figure 2.4C, D). Again, the analyses were difficult because we saw little difference in mobility with the single digests using ODN C (data not shown), B, D or E. However, multiple double digests using these ODNs indicate that no cross-links exist between ODNs A and E. Digestion of gA6-14-

U10 x A6U (B2) with ODN F did generate one fragment of faster mobility (compare with F alone in gA6-14-U5 x A6U, B2), indicating that the cross-link is located downstream of F. In contrast, digestion of gA6-14-U10 x A6P3 with ODN F' generated two fragments suggesting that cross-links did exist both upstream and downstream of F'. Double digests with A+B, B+D, B+E, D+F' and B+F' did not generate any specific fragment (band 3) that could confirm a cross-link between A and F.

In summary, the cross-linking experiments generated two populations of mRNA/gRNA complexes, B1 and B2. Using both reverse transcriptase and RNase H analyses, B1 cross-linked sites were localized to a region upstream of the first editing site. The mapped cross-linked sites for the U5 or U10 modified gA6-14 were complementary, indicating that the U-tail interaction is strongest with nucleotides 25–32 within both A6U and A6P3. The identification of multiple cross-links within a small range of nucleotides suggests a flexible interaction of the U-tail with the mRNA in this region. In addition, for a population of molecules, the U-tail interaction was farther upstream with cross-links detected at nucleotides 7-8. For both A6U and A6P3, the cross-linking patterns were almost identical, indicating that the U-tail is interacting with the same region. In contrast, the B2 cross-link sites were difficult to map by RT, and RNase H analyses suggest that most of the cross-links were downstream of the ABS. A small population of the gA6-14-U5 modified guide RNA did cross-link to nucleotides located near the 5'-end of the mRNA. These upstream cross-links may be a product of 5'-end mRNA folding close to the U-tail. The mapped B2 sites downstream from the ABS indicate potential tertiary folding of the U-tail and anchor helices close to each other.

## **Solution Structure Probing**

Enzymatic and chemical structure probing in solution are used to study the existence and strength of RNA structures under nearly physiological conditions (Cech et al. 1983; Ehresmann et al. 1987; Knapp 1989; Leung and Koslowsky 2001a). In mRNA/gRNA bimolecular structures, the anchor helix, is presumed to form due to the complementarity of the mRNA ABS, immediately 3' of the editing site, and gRNA anchor, at the gRNA's 5'-end (Blum et al. 1990). However, the weak binding affinity of some mRNA/gRNA pairs allows for the presence of unpaired RNAs, preventing accurate structure probing. To counteract this problem, the use of cross-linked RNAs facilitated base pairing of the anchor region and assured folding of the two molecules together (Leung and Koslowsky 2001a; Yu and Koslowsky 2006). 5'-Cross-linked RNAs have been shown to support editosome assembly (Leung and Koslowsky 2001a). Incubation of cross-linked substrates in editing active mitochondrial fractions results in accurate gRNA-directed cleavage. In addition, at sites of U-deletion (the cross-linked A6U and A6P3 substrates), U-specific exonuclease activity on the 5' cleavage products was observed (Leung and Koslowsky 2001a). This indicates that the editing complex interacts and assembles correctly on the cross-linked RNAs and that structures determined utilizing these molecules are biologically relevant. Determination of the affinity constant for the A6 pair indicated that gA6-14 has a high affinity for its target mRNA (Koslowsky et al. 2004) suggesting that cross-linking of the two molecules might not be necessary. In this study, we conducted solution structure probing experiments using both cross-linked and non-cross-linked A6 RNA complexes. Direct comparisons showed no differences in the digestion patterns, indicating that simple hybridization was sufficient for generating complexes stable enough for solution structure probing. Eliminating the cross-linking

step also allowed us to probe the RNA structures using both 5'- and 3'-end labeled molecules. In addition, to facilitate the probing, the mRNAs were shortened at their 5'-ends (sequence that does not interact with the U-tail), based on our U-tail cross-linking results described above. Again, shortening of the A6 substrate at its 5'-end has no effect on its ability to undergo editing when incubated with glycerol gradient purified editosomes (Burgess et al. 1999; Cruz-Reyes et al. 2001; Lawson et al. 2001; Igo et al. 2002; Cifuentes-Rojas et al. 2005; data not shown).

The solution structure probing experimental approach was similar for all treatments. The RNA was first <sup>32</sup>P-end labeled (either mRNA or gRNA, one at a time, at the 5'- or 3'-end), gel-purified, denatured and then renatured alone or in the presence of its interacting partner under native conditions (10 mM MgCl<sub>2</sub>, 100 mM KCl, 10 mM Tris-HCl, pH 7.5; at 27°C for 3 h). To assess single stranded (ss) regions we utilized RNase T1 (specifically cleaves guanines), RNase T2 and Mung Bean (MB) (preference for adenines) (Knapp 1989), and the chemical diethyl pyrocarbonate (DEPC) to modify the N7 position of unstacked adenines and guanines (Peattie 1979; Ehresmann et al. 1987). Double stranded (ds) or stacked areas were identified with cobra venom RNase V1 (Lockard and Kumar 1981). Representative examples of gels used to separate the digestion products are shown in Figure 2.5.

*gRNA structure (gA6-14, gA6-14sU and gA6-14-C6) (Figs. 2.5B, 2.6A)*

The secondary structure for gA6-14 and four other gRNAs has been previously determined and proposed to be a single or a double stem-loop separated by single stranded nucleotides with a single stranded U-tail (Schmid et al. 1995; Hermann et al. 1997; Golden and Hajduk 2006; Yu and Koslowsky 2006). The finding of a common

conformation amongst gRNAs of distinct sequences was suggestive of a binding site for specific proteins that could facilitate editosome assembly (Muller and Goring 2002; Schumacher et al. 2006). Our goal was to find the secondary structure of gA6-14 in a binary complex with the mRNA and how the U-tail influences the structure. As a control, we additionally obtained the gA6-14 structure in solution. While our conditions and sequence were slightly different from Schmid et al. (1995), we also observed two stem-loops in gA6-14. Accessibility to single-stranded specific enzymes within nucleotides 6-8 and to the double stranded specific V1 enzyme on nucleotides 11-13 defined a small stem-loop, which we call stem-loop I (SL I), present within all three gRNA constructs. SL I included the 5'-end of the anchor within the loop. Stem-loop II was larger, formed by the guiding nucleotides, and persisted even after ablation of the U-tail, in gA6-14sU. Digestion by the V1 enzyme and protection against single stranded specific probes at nucleotides 23-29 and 38-45 defined the stem of SL II within gA6-14 and gA6-14sU. The apical loop of SL II was highly accessible to single stranded specific probes, confirming their unpaired nature. The substitution of 6 U's from the U-tail by 6 C's in gA6-14-C6 changed the digestion/modification pattern within SL II, forming two alternate stem-loops (stem-loop IIa and b, Fig. 2.6A). SL IIa involved 3 base pairs (bp) (nucleotides 14-16 and 22-24) confirmed by V1 digestions at nucleotides 14, 23 and 24 and mfold predictions (Walter et al. 1994; Zuker 2003). The larger SL IIb involved the C6-tail and was confirmed by the presence of V1 digestion within both sides of the stem (nucleotides 32-44 and 48-60). Weak accessibility by single stranded specific enzymes and DEPC within the stem is explained by the presence of several single-nucleotide

mismatches that destabilize the helix. Nonetheless, this gA6-14-C6 seems to be thermodynamically more stable than the parental gRNA .

*mRNA structure (A6U and A6P3) (Figs. 2.5A, 2.6A)*

The 3'-end of the unedited A6 mRNA is purine-rich, thus predicted to be mostly unstructured (Koslowsky et al. 2004). Solution structure probing of A6U alone (short and long constructs) indicated that its most stable structure involves the formation of a single (8-10 bp) stem with an 11-nt loop. V1 cleavages within a few loop nucleotides suggest that they may interact or stack. Note that the ABS is mostly contained within the terminal loop. When the mRNA is partially edited through three sites (A6P3) the 11-nt terminal loop was maintained. However, the sequence changes associated with RNA editing introduced new base pairings within the stem region.

**Figure 2.5. Solution Structure Probing. (A)** 5'-end labeled A6U and A6P3 alone or hybridized to gA6-14 (+14), gA6-14sU (+sU), and gA6-14-C6 (+C6). **(B)** 5'-end labeled gA6-14, gA6-14sU, and gA6-14-C6 alone and hybridized to A6U and A6P3. (Anchor) sequence complementary to A6U-ABS. (Extension) anchor sequence extended, complementary to A6P3-ABS. Figures are representative of denaturing polyacrylamide gels from partial digestions of RNAs with RNases T2 (single stranded A specific) and V1 (specific to double stranded or stacked nucleotides) or DEPC (chemical specific to non-stacked nucleotides). (T1) ladder lane produced by partial digestion under denaturing conditions with RNase T1. (NE, NC) no enzyme or chemical controls. (Numbers above each lane) digestion time in minutes. (Numbers next to bands) nucleotide position according to Figure 2.1.

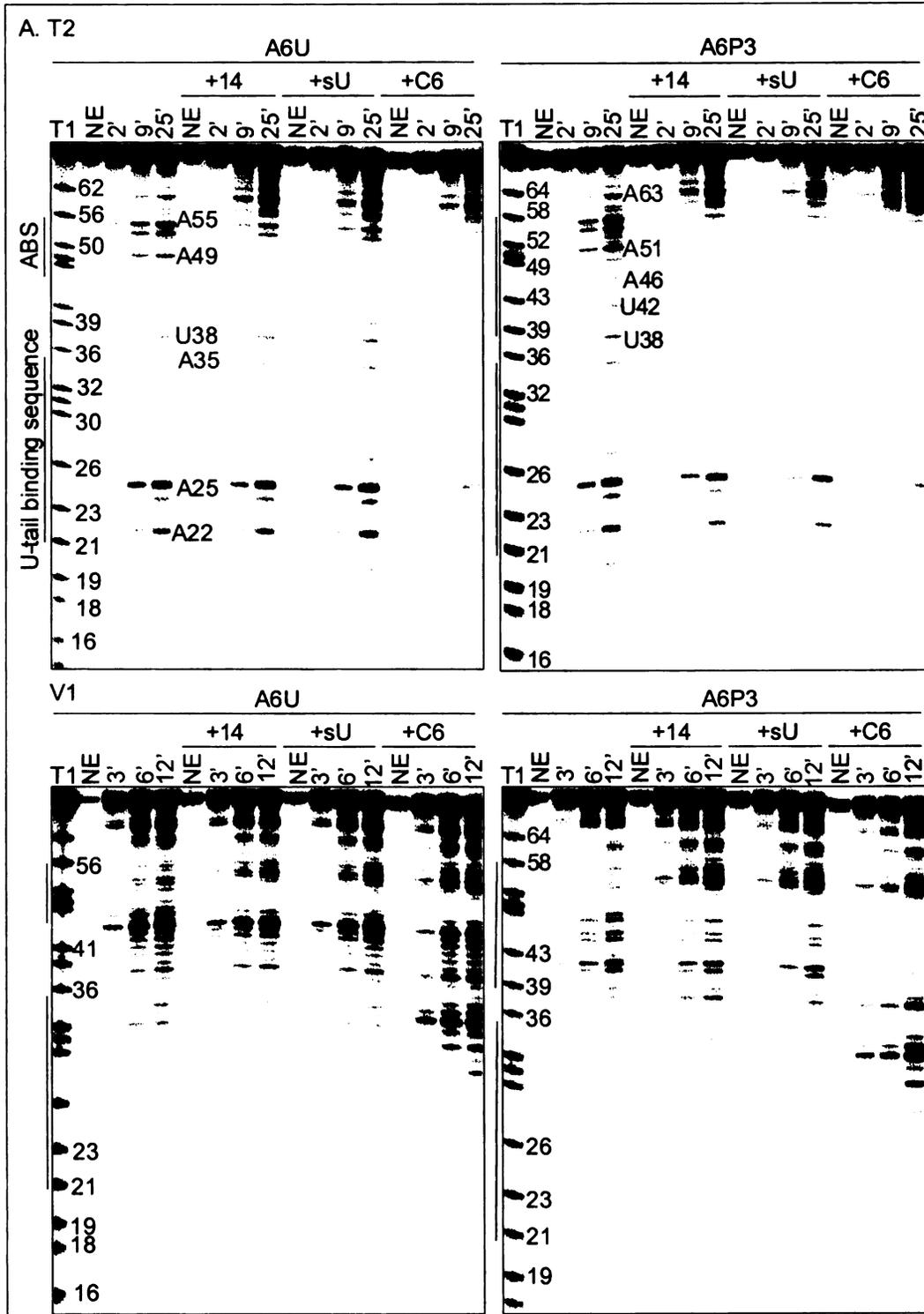


Figure 2.5 Solution Structure Probing. 5'-end labeled A6U and A6P3 alone or hybridized to gA6-14 (+14), gA6-14sU (+sU), and gA6-14-C6 (+C6).

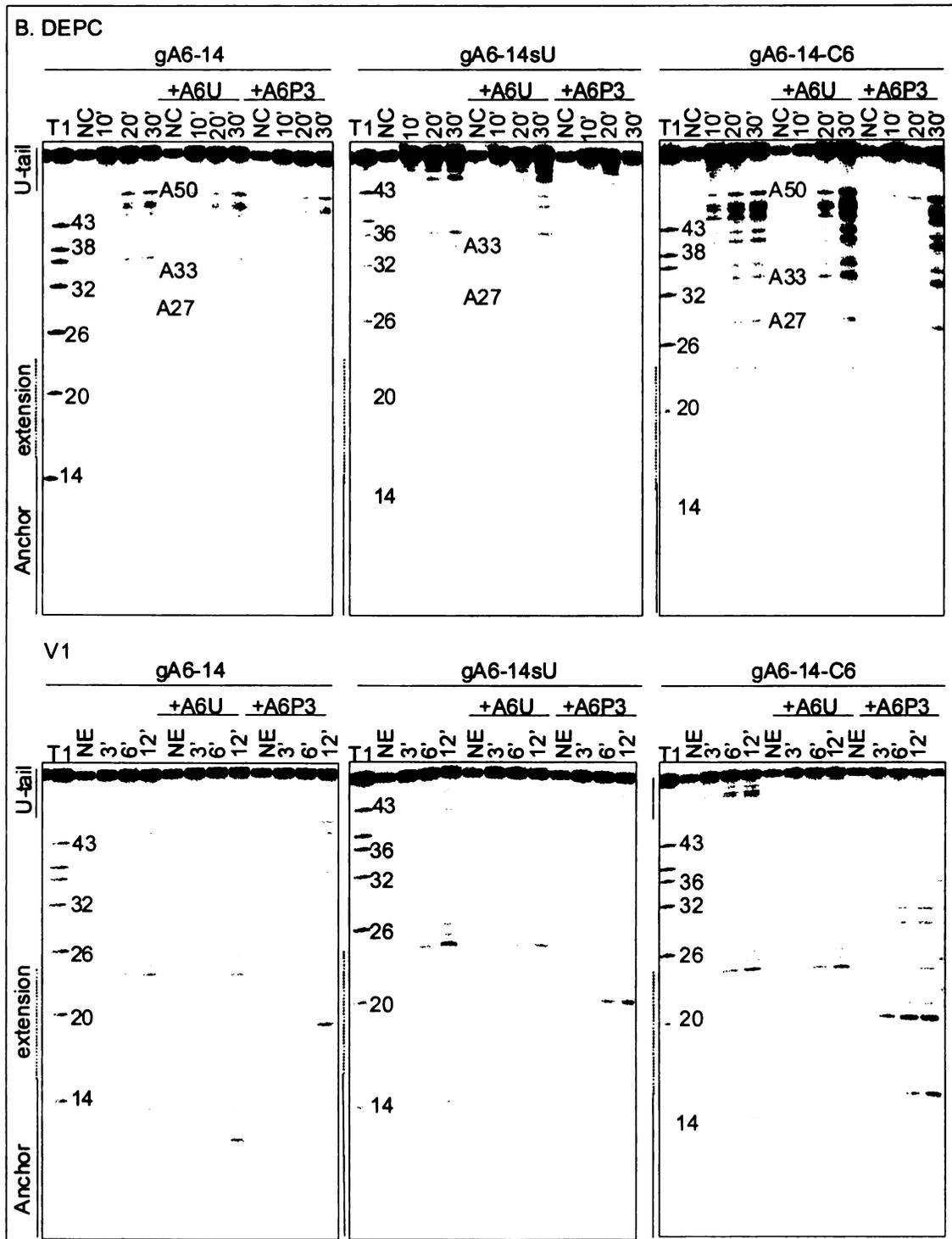
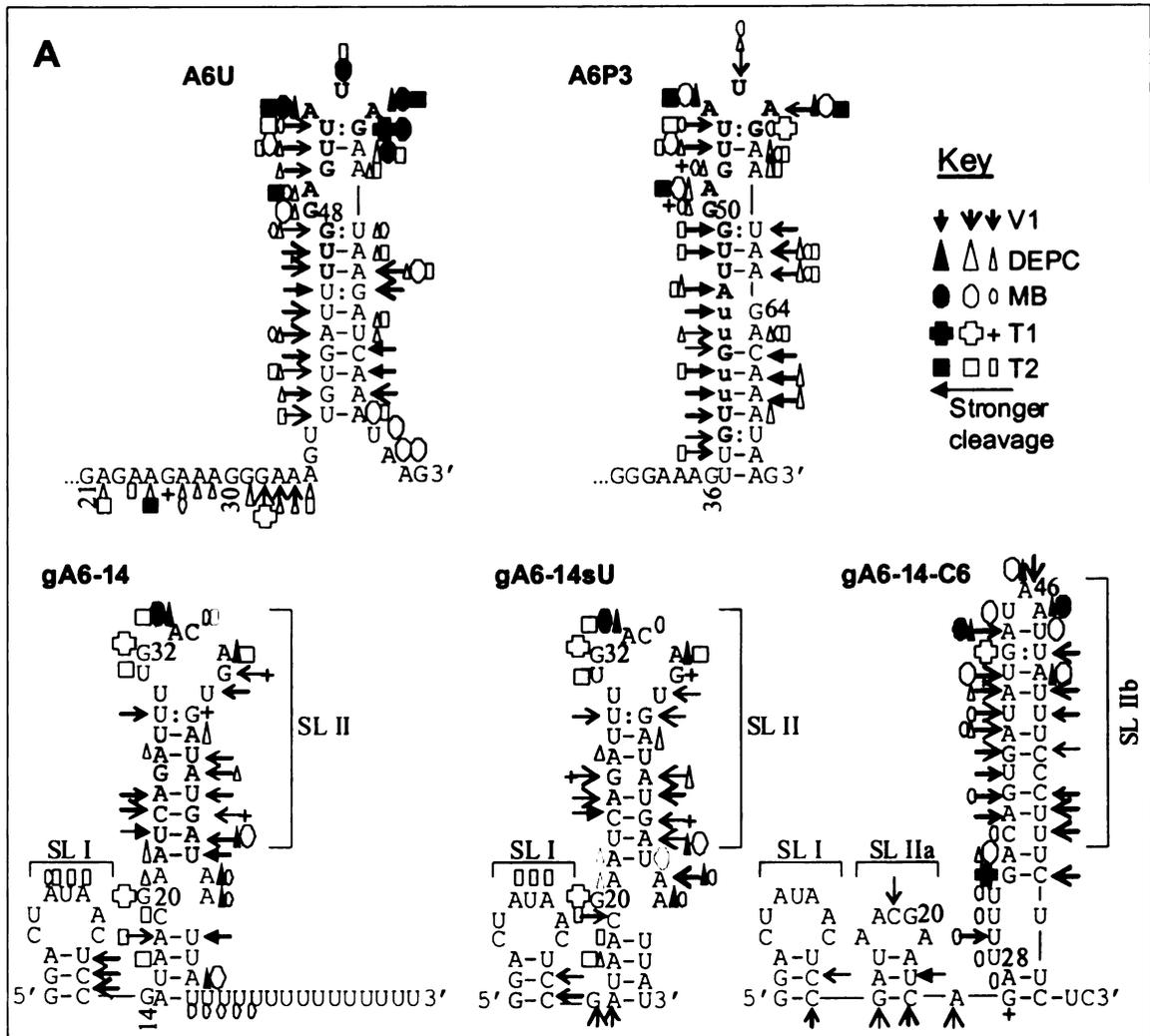


Figure 2.5 (cont'd).

*mRNA/gRNA structure (Figs. 2.5, 2.6B)*

The secondary structure in solution for 6 mRNA/gRNA pairs (A6U and A6P3 with gA6-14, gA6-14sU, and gA6-14-C6) is described below, by motif, starting with the anchor helix, then the U-tail helix, the gRNA stem-loop, and lastly the mRNA junction at the intersection of the three helices.

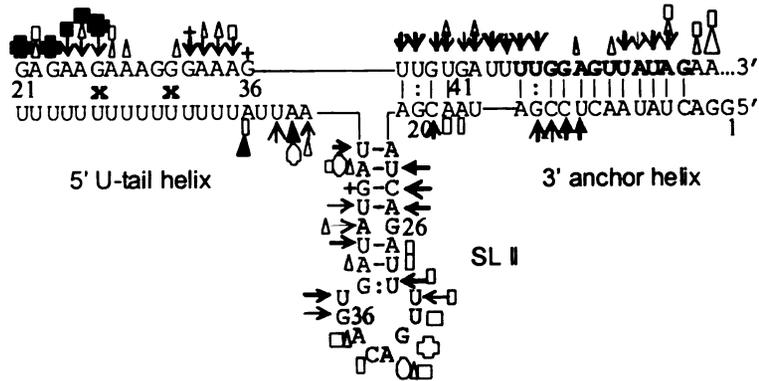
Anchor helix-A6U and A6P3 hybridized to gA6-14, gA6-14sU, and gA6-14-C6. The anchor helix within A6U/gA6-14 involves formation of 12 contiguous base pairs formed between the ABS of the mRNA and the anchor of the gRNA. Structure probing after cross-linking or hybridization indicates that both the ABS and the gRNA anchor are protected from single-stranded nucleases, confirming proper binding and formation of the recognition helix. When the mRNA was not cross-linked to the gRNA, its 3'-end was more sensitive to single stranded probes as opposed to the cross-linked molecules. This higher sensitivity was expected and due to termini breathing because the molecules were not covalently linked. The sensitivity was more noticeable and extended throughout the ABS when A6U was paired with gA6-14sU, suggesting that the presence of a gRNA U-tail increases the stability of the 3' anchor helix. Anchor helix formation was confirmed by RNase V1, where V1 dependent cleavages were observed at both ends of the duplex. V1 digestion at the helix ends was expected considering that V1 does not digest the entire extension of a duplex (Lockard and Kumar 1981). V1 requires a minimum of 4-6 sugar phosphate residues within a helix to bind and cleave the phosphates on either strand (Lowman and Draper 1986). A similar V1 digestion pattern, cleaving at both ends of a helix, has been previously reported for the CYb anchor helix (Yu and Koslowsky 2006).



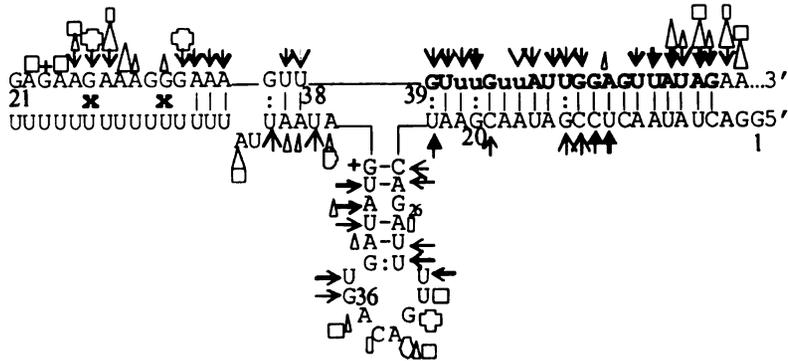
**Figure 2.6** Secondary structure models for the individual A6 RNAs and mRNA/gRNAs during distinct stages of editing. **(A)** Individual RNAs, A6U, A6P3, gA6-14, gA6-14sU, and gA6-14-C6. The ABS within the mRNAs and the anchor region within the gRNAs are bolded. The anchor extends to nucleotide 23 when the gRNAs are paired with A6P3. Stem-loops (SL) are indicated next to gRNAs. In gray is SL II, conserved in many gRNA species. **(B)** mRNA/gRNA complexes, as indicated: A6U and A6P3 with gA6-14, gA6-14sU, and gA6-14-C6. The structures were obtained with a combination of data obtained from cross-linking analysis (x) (mRNA/gRNA pairs only), computer algorithm, and solution structure probing using five different enzymes or chemical. RNase V1 (→) has specificity to double-stranded or stacked nucleotides, T1 (+) is specific for single stranded G, T2 (□) and Mung Bean (○) are mostly specific for single stranded A's, but can also modify other nucleotides. The chemical DEPC (Δ) is specific for unstacked purines. Symbol size indicates intensity of cleavage and filled symbols indicate the strongest cleavage sites.

**B**

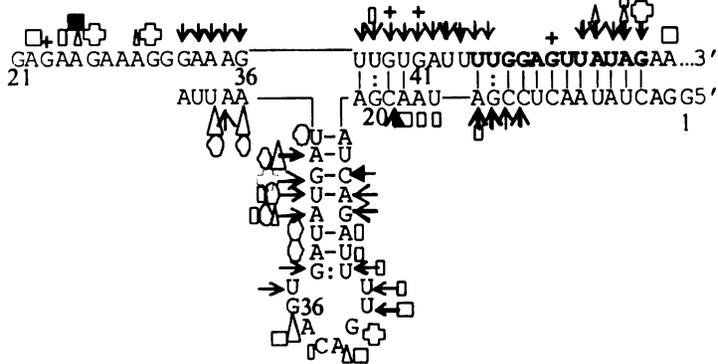
**A6U/gA6-14**



**A6P3/gA6-14**



**A6U/gA6-14sU**



**A6P3/gA6-14sU**

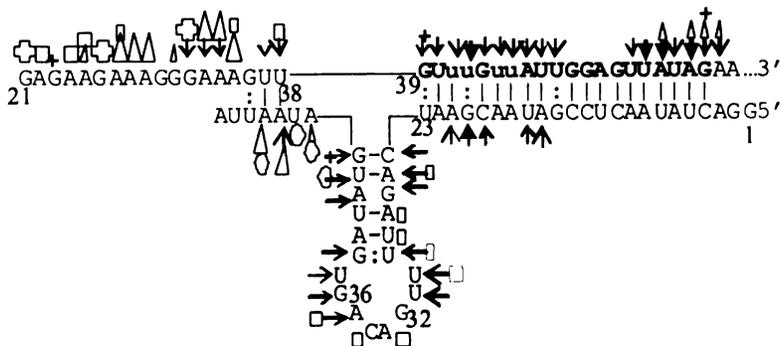


Figure 2.6 (cont'd).



Editing of the first 3 sites in A6P3 extends the anchor helix by 8 additional base pairs. Similar to A6U, the 3'-end of A6P3 ABS was lightly cleaved by single-strand specific probes when the complex was not cross-linked. Nonetheless, both sides of the helix were protected from cleavage by ss specific probes. In contrast to A6U, an increased sensitivity to the ss specific probes was not observed when A6P3 was paired with gA6-14sU, reflecting the increased stability introduced by the additional base pairs. Again, the V1 digestion pattern was complementary to the other probes.

U-tail helix-A6U and A6P3 hybridized to gA6-14 and gA6-14-C6. UV cross-linking experiments indicate that A6/gA6-14 complexes form a U-tail helix involving the gRNA U-tail and the purine region G21 to A35 within A6U and A6P3. While the strongest cross-links were mapped to nucleotides 29-31 for U5 and 26-28 for U10, a number of cross-links were identified suggesting that the U-tail interaction is dynamic within this region. Solution structure probing indicated that most of this region was susceptible to single strand specific probes, and the few V1 cleavages were inconsistent and light. This indicates that the U-tail forms a relatively unstable, flexible duplex with the purine-rich region. Duplex instability has been associated with G:U wobble base pairs (Freier et al. 1986) which are unlikely to stack (Auron et al. 1982) and therefore are not recognized by RNase V1. The enhanced gA6-14-C6 introduces 6 G:C base pairs that strengthen binding and this allowed structure probing of the helix. Decreasing the number of G:U base pairs and increasing the G:C content reduced the number of sites accessible to single-stranded probes and increased V1 products, confirming stable helix formation.

Stabilizing the U-tail interaction also seemed to stabilize the gRNA stem-loop, described below.

gRNA stem-loop II (SL II)-A6U and A6P3 hybridized to gA6-14, gA6-14sU, and gA6-14-C6. As described above, gA6-14 alone forms a stable stem-loop (SL II) that involves the guiding region nucleotides (Fig. 2.6A). Similar to the cytochrome b interaction with gCYb-558 (Yu and Koslowsky 2006) this stem-loop appears to be maintained during the initial interactions with the mRNA. V1 digestions from U23 to U29, and between G38-U45 (bases within the stem) were detected in gA6-14 alone as well as when the gRNA was cross-linked or hybridized to both A6U and A6P3. The apical loop within gA6-14, formed with both mRNA pairings, was readily accessible to single stranded probes. RNase V1 lightly cleaved at the loop-nucleotides (U30-31 and G36-U37) near the closing nucleotides. This V1 pattern of loop digestion has been previously observed and can be caused by a number of conditions, which include over-cutting the terminal ends of a double stranded region (Auron et al. 1982), higher order structure in the loop, like stacking of loop nucleotides with the stem (Krol et al. 1990), Watson-Crick or mismatch hydrogen-bonds between loop bases (Burkard et al. 1999), or tertiary interactions (Lockard and Kumar 1981). Editing of A6U through 3 editing sites (A6P3) does disrupt the first two base pairs of the gRNA stem-loop II. However, the loss of these two base pairs did not destabilize the rest of the gRNA stem-loop structure.

Light cleavage by single stranded probes was observed within the gRNA stem region, most likely due to some destabilization caused by the single A:G mismatch (Figs. 2.5B, 2.6B). However, the stability of the stem also appears to be influenced by the

presence of the upstream U-tail helix. Deletion of the U-tail (gA6-14sU) causes an increase susceptibility of the 3' side of the stem to ss specific probes. In addition, strengthening of the U-tail interaction by using the modified gA6-14-C6 decreased the intensity and the number of sites available for the ss specific probes. This suggests that the U-tail not only stabilizes the anchor helix but also the gRNA stem-loop II. For the A6P3 pairing, RNase V1 cleavages were observed at the junction between the anchor helix and gRNA stem-loop, suggesting continuity between both helices. Continuous V1 digestion at RNA junctions, indicating helical stacking, has been described in other RNAs such as tRNAs, U1 snRNA, and cucumber mosaic virus satellite RNA (Auron et al. 1982; Lowman and Draper 1986; Krol et al. 1990; Bernal and Garcia-Arenal 1997; Lescoute and Westhof 2006). Although this V1 pattern was not observed in the A6U/gA6-14 pair, we did observe an absence of single strand specific probe accessibility, which is also suggestive of continuous double-stranded stacking of two helices.

Junction-A6U hybridized to gA6-14, gA6-14sU, and gA6-14-C6. The connecting sequence mRNA 5'-UUGUG<sub>41</sub>AUU<sub>44</sub>-3', partially paired with gRNA 3'-AG<sub>20</sub>CAAU<sub>16</sub>-5', forms an asymmetrical internal loop at the 5'-end (mRNA orientation) of the anchor helix, next to the SL II and the U-tail helix (Figs. 2.5, 2.6B). The models suggest that there are nucleotides within this asymmetrical loop that are not base paired; however, when A6U/gA6-14 was analyzed, these junction nucleotides were not digested by ss specific probes, but were susceptible to RNase V1. V1 cleaved the 5'-end of the mRNA ABS, extended through the three first junction nucleotides, and cleaved again at the 4 bp stem. While this V1 digestion pattern was maintained when A6U was folded in the

presence of gA6-14sU and gA6-14-C6, the mRNA nucleotides at the junction appeared to be more available to ss specific probes. The strongest V1 cleavages were at the bonds between nucleotides 42-44 within the 4x2 internal loop, indicating that these loop nucleotides could be interacting with the opposing gRNA nucleotides or stacking between the two stems. On the gRNA, V1 cleaved at C19-G20, confirming the 4 bp-stem formation. V1 only moderately distinguishes stacked single stranded from base paired nucleotides (Lowman and Draper 1986). Therefore, it is likely that these junction nucleotides follow the same conformation present in the anchor helix. Analyses using gA6-14-C6, which locks the U-tail in its position, does not change the cleavage patterns at the junction.

Junction-A6P3 hybridized to gA6-14, gA6-14sU, and gA6-14-C6. The solution structure probing of the junction regions between the anchor helix, SL II, and the U-tail helix for the gRNA interaction with A6P3 are shown in Figures 2.5 and 2.6B. This mRNA region (nucleotides 36-38) was not susceptible to cleavage by single stranded probes, and similar to the unedited mRNA, was cleaved by RNase V1. Again, this suggests that the nucleotides surrounding the next available editing site are in a helical or stacked conformation. Small differences were detected upon deletion of the gRNA U-tail, such as, the V1 cleavage sites within the anchor and more intense single stranded specific digestions at the gRNA 3'-end (U<sub>48</sub>-A<sub>42</sub>). This single strand specific digestion in gA6-14sU disappeared when A6P3 was hybridized with the enhanced gA6-14-C6, reinforcing the finding that the U-tail promotes stabilization of the U-tail helix and the gRNA stem-loop.

## **DISCUSSION**

RNA structure promotes editing site specificity in several systems (Connell and Simpson 1998; Koslowsky 2004; Reenan 2005). In kinetoplastids, natural editing sites are flanked by an upstream interaction between a mRNA purine-rich region and the gRNA U-tail, the U-tail helix (Blum et al. 1990; Kable et al. 1996; Leung and Koslowsky 1999) and by a downstream anchor helix (Blum et al. 1990; Byrne et al. 1996; Cruz-Reyes and Sollner-Webb 1996; Kable et al. 1996; Seiwert et al. 1996; Adler and Hajduk 1997). Our results indicate that a third helix, formed by the gRNA, is also present at the junction. We hypothesize that the three helices are part of a core structure that exists in most natural mRNA/gRNA complexes (Leung and Koslowsky 1999; 2001a; 2001b; Yu and Koslowsky 2006). This core three-helical structure has been confirmed by solution structure probing of the A6/gA6-14 complex (this paper) and the CYb/gCYb-558 complex (Leung and Koslowsky 2001a; Yu and Koslowsky 2006).

In this work, we propose secondary structure models for six A6 mRNA/gRNA pairs obtained by a combination of point constraints coupled with computer folding of primary sequence, plus nuclease and chemical probing. Independently of being unedited (A6U) or partially edited through the first 3 editing sites (A6P3), the secondary structure for the A6 mRNA/gRNA complex consists of three helices.

The anchor helix is a perfect 12 base-paired duplex in A6U/gA6-14 and is extended an additional 8 bp in A6P3/gA6-14. Although our solution structure probing experiments indicate that the U-tail helix is unstable, the cross-linking data and the use of gA6-14-C6 confirmed the U-tail binding site within A6U and A6P3 to be ~10 and 3 nt upstream of the ES, respectively. gA6-14-C6 is a construct developed by Cruz-Reyes et

al (2001), that strengthens U-tail helix affinity and increases editing efficiency by a factor of 4. Cruz-Reyes et al (2001) observed that U-deletion efficiency in vitro can be further increased when the A6 mRNA/gRNA pair forms relatively stable anchor and U-tail helices and lacks the third helix (gRNA stem-loop II). The most efficient gRNA construct, D33, was shown to be >100 times more efficient than gA6-14- $\Delta$ 16G anchor, with only 3 Cs as guiding nucleotides, and with a modified tail (3'-CCCUUCAAUAU-5') that binds strongly 3 nt upstream the ES. The native gRNAs did not evolve to have D33 or gA6-14-C6 features because they could not direct sequential cycles of editing in vivo and the G:C rich hybrid would be too stable to dissociate. The above work by Cruz-Reyes et al. (2001) and others show that the position and strength of the U-tail helix formation with the mRNA directly influences editing efficiency (Seiwert et al. 1996; Burgess et al. 1999; Kapushoc and Simpson 1999; Igo et al. 2002; Golden and Hajduk 2006; Cifuentes-Rojas et al. 2007). To evaluate the structural changes caused by distinct U-tails, here we used the parental gA6-14, and gA6-14 without its U-tail (gA6-14sU) and with an enhanced U-tail (gA6-14-C6). We selected gA6-14-C6 because it is identical to the parental gA6-14 in sequence except for the C-U substitutions within the U-tail. In fact, gA6-14-C6 generated the same overall bimolecular structure observed with gA6-14, except that the higher stability of the U-tail helix seemed to additionally improve the stability of SL II. The structures generated with gA6-14-C6 support the structures obtained with gA6-14. In contrast, deletion of the U-tail, (gA6-14sU), caused more instability within SL II in both the A6U and A6P3 complexes. In addition, deletion of the U-tail further decreased the stability of the anchor helix in the A6U complex. Thus, in addition to confirming the binding site for the U-tail within the unedited and partially

edited A6 mRNAs, we found that addition of the U-tail helix stabilizes the other two helices.

In vitro, the gRNA stem-loop has been shown to be dispensable and even inhibitory to U-deletion editing at ES1 in A6 (Seiwert et al. 1996; Cruz-Reyes et al. 2001; Cifuentes-Rojas et al. 2005). However, another study indicates that the gRNA stem-loop structure, can be essential for proper editing for some substrates (Golden and Hajduk 2006). In this report, we show that the SL II formed within gA6-14 alone is maintained after gA6-14 annealing with both A6U and A6P3, corroborating earlier computer predictions (Leung and Koslowsky 1999). Locking the gA6-14 U-tail into a single position by using the modified gA6-14-C6 construct did not affect the formation of SL II in either A6U or A6P3 complexes. Maintenance of SL II going from unedited to partially edited A6P3 occurred by simple rearrangements within the guiding region, without displacement of the U-tail. In previous studies with gCYb-558, the gRNA stem-loop was also maintained upon gRNA/mRNA annealing (Yu and Koslowsky 2006). In CYb, however, extension of the anchor helix in the partially edited CYbPES3 disrupts base-pairs within the original stem-loop structure, with a new stem-loop formed by incorporation of nucleotides from the U-tail. The preservation of the original gRNA stem-loop after annealing with the mRNA, observed in both the A6 and CYb complexes, suggests that the formation of a three helical junction is thermodynamically favorable. Furthermore, conservation of the gRNA SL structure through initial stages of editing in both A6 and CYb mRNA/gRNA complexes supports our hypothesis that a three-helical structure is important for stabilizing the gRNA/mRNA interactions during the initial editing events.

The junction nucleotides between the three helices include the first editing site for A6U (determined ES1) and A6P3 (ES4), which are both deletion sites. The results of the structure probing indicate that ES1 and ES4 have similar features. When both mRNAs were hybridized with gA6-14, gA6-14sU and gA6-14-C6, the ES and the sequence immediately upstream (6-8 nt) were inaccessible to single strand specific probes and only accessible to the double-strand- or stacking-specific V1. This suggests that while the scissile base is not involved in Watson-Crick pairing, in agreement with predictions by Seiwert et al. (1996), Cruz-Reyes et al. (2001), and Lawson et al. (2001), it is contained within a highly organized region. This finding was not surprising because nucleotides in single stranded areas in a secondary structure are often paired in long-range tertiary interactions, single strand base stacking, or cross-strand base stacking (Peterson and Feigon 1996; Butcher et al. 1997; Zimmermann et al. 1997; Schroeder et al. 2004; Znosko et al. 2004). Internal loops or bulges are common in RNA structure and create distortions into the geometry of an RNA helix that are often critical for protein recognition (Burkard et al. 1999). For instance, internal loops within 5S rRNA have a closed conformation due to several noncanonical interactions that are well suited for protein-binding (Westhof et al. 1989; Brunel et al. 1991). Furthermore, the editosome has been shown to directly interact with ES1 and ES2 within the A6 mRNA/D33 complex (Sacharidou et al. 2006; Cifuentes-Rojas et al. 2007). Consistent with this notion, the structural organization observed near A6 ES1 and ES4 in our study may be a higher-order determinant of editosome binding. Analyses of the unedited CYbU mRNA complexes with gCYb-558 indicate that the nucleotides near ES1 are highly accessible to single strand specific probes. Editing in the CYb mRNA is different from A6 because it

involves only uridylyte insertions. In contrast, editing in A6 begins with a deletional event followed by insertional editing at sites 2 and 3, with site 4 requiring another deletional event. Small structural differences surrounding the immediate editing site, like the ones found in this study, may be important for recognition by the correct editosome subcomplex, as suggested previously (Schnauffer et al. 2003; Panigrahi et al. 2006; Panigrahi et al. 2007). Differences in the efficiency of full-round editing assays have also been found according to the type of ES within the same mRNA and between distinct mRNAs, supporting the idea that small structural differences can influence editosome assembly (Cifuentes-Rojas et al. 2005).

In summary, structure probing of the A6/gA6-14 editing pair indicates that the overall structure formed between the initiating gRNA and its cognate mRNA involves the formation of three helices that interact to stabilize the RNA bimolecular complex. Similar to the previously studied CYb/gCYb-558, the gRNA stem-loop is maintained in its initial interaction with the mRNA, and the U-tail forms a helix with purine-rich mRNA sequences just upstream of the first few editing sites. Formation of the U-tail helix helps stabilize both the gRNA stem-loop and the anchor helix. In addition, we show that the gA6-14 stem-loop is also maintained during the initial editing events. The overall secondary structure for the A6/gA6-14 complex was very similar to that observed for the CYb/gCYb558 interaction, except for peculiar differences in the structure surrounding the immediate editing site. While the CYbU insertional site was found to be open and accessible to ssRNA probes, both of the A6 deletional sites were found to be within highly organized regions.

## MATERIALS AND METHODS

### Oligodeoxynucleotides (ODN)

All ODNs (Table 1) were chemically synthesized and obtained from Integrated DNA

Technologies, Inc.

Table 2.1 List of oligodeoxynucleotides

ODN	Sequence (5' to 3')	Length (nt)
gA6-14sU	TAATTATCATATCACTGTCAAATCTGATTCGTTATCGGAGTTATAGC CCTATAGTGAGTCGTATTAATT	71
gA6-14	AAAAAAAAAAAAAAAAATAATTATCATATCACTGTCAAATCTGATTCG TTATCGGAGTTATAGCCCTATAGTGAGTCGTATTAATT	86
T7-22	AATTTAATACGACTCACTATAG	22
bridge g14	AAAAAAAAAAAAAAAAATAATTATCATAT	24
bridge gC6	GAGAAGAAAGGGAAATAATTATCATATCAC	30
RT-1	TATTATTAACCTATTTGATCTTATTCTATAAC	32
RT-2	TATTATTAACCTATTTGATC	20
RT-3	TATTATTAACCTATTTG	17
RT-4A6U	CTCCAAAATCACAACCTTC	19
RT-5A6P3	CTCCAATAACAAACAACCTTC	21
A	GTTTGGCAA	10
B	TCCTGCTCTT	10
C	CCTAACCTTTCCTGC	15
D	TCCTCCCCCT	10
E	TTCTTCTCTCCTCCCCCT	18
F	TCACAACCTT	10
F'	CAAACAACCT	10
G	CTATAACTCC	10

### Oligoribonucleotides

The oligoribonucleotides below were chemically synthesized and obtained from

Dharmacon Research, Inc. The C6-tail was obtained from Integrated DNA technologies.

U<sub>15</sub>-tail (15 nt): 5'-UUUUUUUUUUUUUUU-3'

C6-tail (15 nt): 5'-UUUCCCUUCUUCUC-3'

U5-tail (15 nt): 5'-UUUU-s<sup>4</sup>U-UUUUUUUUUU-3'

U10-tail (15 nt): 5'-UUUUUUUUU-s<sup>4</sup>U-UUUUU-3'

### DNA Templates and RNA synthesis and modification

The procedures below have been previously described (Leung and Koslowsky 2001a;

Koslowsky et al. 2004; Yu and Koslowsky 2006). DNA templates were PCR

amplifications from plasmids. The mRNAs were synthesized either by T7 RNA polymerase (RiboMax, Promega) according to the manufacturer's directions or in the presence of 5 mM guanosine 5'-monophosphorothioate (GMPS). Templates for gRNA synthesis were prepared by hybridizing the T7-22 ODN to gA6-14sU and gA6-14 ODNs. The transcription of the gRNAs was based on the Uhlenbeck single-stranded T7 transcription using T7 RNA polymerase (Ambion, Inc.) (Milligan et al. 1987; Milligan and Uhlenbeck 1989). The gRNAs were synthesized in the presence (Burgin and Pace 1990; Harris and Christian 1999) or in the absence of 10 mM GMPS (Biolog Life Science Institutes or Dr. Michael E. Harris laboratory at Case Western Reserve University School of Medicine, Cleveland, Ohio). In addition these transcripts also contained traces of [ $\alpha$ - $^{32}$ P] rATP (Perkin Elmer) for visualization and recovery. Half of the GMPS-gA6-14sU transcripts were ligated to a 5'-end labeled U<sub>15</sub>-tail. Half of the sample containing the thiol group at the 5'-end (GMPS-gRNA) was then coupled to azidophenacyl bromide. End labeling was performed with T4 Polynucleotide Kinase (Invitrogen) and 300  $\mu$ Ci of [ $\gamma$ - $^{32}$ P] ATP for 1 nmol of U<sub>15</sub>, U5, or U10-tail, 50  $\mu$ Ci for 25 pmols of free mRNA or 10 pmols of cross-linked mRNA or for 20 pmols of ODN for primer extension analyses. 3'-End labeling involved ligation of 130  $\mu$ Ci (43 pmols) of cytidine 3',5'-biphosphate (pCp) to about 500 pmols of GMPS-gA6-14sU and GMPS-gA6-14 with T4 RNA ligase (New England Biolabs) as per manufacturer's directions. RNA products were separated by electrophoresis on 8% (w/v) (mRNA) or 15% (w/v) (gRNA) polyacrylamide gels containing 8M urea and eluted overnight in elution buffer (10 mM Tris, pH 7.8, 0.1% SDS, 2 mM EDTA, and 0.3 M NaOAc, pH 7.0) in the

presence of phenol. The RNA concentrations were determined by measuring the absorbance at 260 nm, using a Cary 50 spectrophotometer.

### **RNA Crosslinking**

Reactions involved the cross-link of gRNA to the mRNA as described before (Leung and Koslowsky 1999). The gRNA contained either 4-thio-U ( $S^4U$ ) or APA groups. Briefly, 50-200 pmol of gRNA was annealed to mRNA (3 molar excess) in HE buffer (25 mM Hepes, pH 7.5, 2 mM MgOAc, 50 mM KCl, 0.5 mM DTT, 0.1 mM EDTA) by heating to 50°C, cooling 3°C/min to 27°C, and incubating at 27°C for 45 min. The hybridized RNAs were placed on ice and irradiated at a distance of 5 cm from Stratalinker (Stratagene) 312 nm bulbs, for 20 min. Cross-linked molecules were ethanol precipitated and gel extracted. The efficiency of cross-links was obtained by exposing the gel to phosphorscreen and using a Storm PhosphorImager (Molecular Dynamics). The cross-links were further mapped and confirmed through primer extension and RNase H analysis, respectively. Alternatively, the cross-linked mRNA was 5'-end labeled using T4 kinase to proceed with solution structure probing.

### **Primer Extension Analysis**

Primer extension procedure has been described before (Leung and Koslowsky 1999). Briefly, 50 Kcpm of a 5'-end labeled primer (we used 5 different RT ODNs) is hybridized to the cross-linked RNAs or 2-5 ng of control mRNA in RT buffer (50 mM KCl, 20 mM Tris-HCl pH 8.5, 0.5 mM EDTA and 8 mM  $MgCl_2$ ). Then, an extension cocktail containing 0.25 U of AMV reverse transcriptase (Seikagaku) and all four dNTPs (1.6 mM each) is added. Primer extension is carried out at 50°C for 30 min and stopped by addition of formamide loading buffer (95% formamide, 20 mM EDTA, 0.05%

bromophenol blue, and 0.05% xylene cyanol). Sequencing reactions required 0.8 mM of each dNTP and 0.4 mM of each ddNTP per reaction. Reactions were resolved on 8-10% (w/v) denaturing polyacrylamide gels that were exposed to a phosphor screen and analyzed using a Storm PhosphorImager.

### **RNase H assays**

The assay has been described before (Leung and Koslowsky 1999). Briefly, 20 pmols of ODN (A, B, C, D, E, F, F', or G) are incubated with ~1000 cpms of the cross-linked molecules, 20 pmols of non-labeled mRNA of the same species, and 2 U of RNase H (Takara). The digestion is held at 55°C for 30 min, under 40 mM Tris-HCl pH7.5, 100 mM KCl, 2 mM MgCl<sub>2</sub>, 1 mM DTT and stopped by addition of formamide loading buffer. Reactions were resolved on 6% (w/v) denaturing polyacrylamide gels that were exposed to a phosphor screen and analyzed using a Storm PhosphorImager.

### **Secondary Structure Prediction**

The mfold computer program (Walter et al. 1994; Zuker 2003) was used to predict the secondary structure of single molecule RNAs used in the present study. The bimolecular secondary structure (mRNA/gRNA) was obtained by RNAstructure version 4.5 (Mathews et al. 2004) or mfold (entering the mRNA and gRNA sequences joined by a linker of 3 non-base pairing N residues).

### **Solution Structure Probing**

Experiments were performed using non-cross-linked or cross-linked mRNA/gRNA pairs (5'-end of gRNA cross-linked to 3'-end of mRNA ABS).

### *Enzymatic probing*

The labeled cross-link or non-cross-linked RNA (200 Kcpms) and the unlabeled partner (0.5 to 1.5 pmols) were hybridized in structure probing buffer (SPB, 10 mM Tris-HCl, pH 7.5, 100 mM KCl, 10 mM MgCl<sub>2</sub>) by heating for 3 min at 70°C, or at 55°C if cross-linked, slowly cooling down to 27°C, and keeping at 27°C for 30 min (if cross-linked) or 3 hours (if non-cross-linked). After the no enzyme (NE) control aliquot was taken, 10 µg of yeast tRNA was added with one of the following enzymes: 0.1 U RNase T1 (Industrial Research, Ltd.), 0.3 U RNase T2 (Invitrogen), 0.25 U Mung Bean Nuclease (MB New England Biolabs) or 0.1 U V1 (cobra venom, Pierce M.B.). RNA digestion was conducted at 27°C, followed by withdrawal of 10 or 20 µl aliquots at 2 or 3 time points. The reaction was stopped at 4°C by addition of TE (10 mM Tris-HCl, pH 7.0, 0.1 mM EDTA) and phenol, followed by phenol/chloroform/isoamyl alcohol (25:24:1) extraction and ethanol precipitation.

### *Chemical probing*

Diethylpyrocarbonate (DEPC, Aldrich) carbethoxylates the atom N7 of adenosines, at neutral pH, that are not involved in helices. This modification opens the imidazole ring between atoms N7 and C8, creating a site for aniline strand scission. The end product is N7-CO<sub>2</sub>H<sub>2</sub>, which runs slower on a gel, compared to cleavages performed by RNases T1 and T2. The reactions were set up as for enzymatic probing as above, then ~6 µl of DEPC were added and incubated at 27°C for different time points. Reactions were stopped by addition of TE buffer and ethanol precipitated twice in 0.3 M sodium acetate (pH 6). Pellets were suspended in 20 µl of 1 M aniline (Aldrich), pH 4.5. Strand scission

was performed at 55°C for 20 min, in the dark and stopped by double ethanol precipitation. The above treated RNAs were suspended in formamide loading buffer and analyzed on 12 or 15% (w/v) polyacrylamide gels containing urea. RNA ladders (20 µl) were generated by enzymatic digestion of heat denatured RNAs at 55°C for 8 min in buffer I (33 mM sodium citrate, pH 5.0; 1.7 mM EDTA, 7 M urea, 0.04% xylene cyanol, 0.08% bromophenol blue) when digesting with T1 (0.028 U) or MB (2 U) and buffer II (33 mM sodium citrate, pH 3.5; 1.7 mM EDTA, 7 M urea, 0.04% xylene cyanol, 0.08% bromophenol blue) when digesting with U2 (0.1U) or T2 (0.0004U). Gels were exposed to a phosphor screen and analyzed by phosphorimager using ImageQuant software from Storm PhosphorImager (Molecular Dynamics).

### ***ACKNOWLEDGMENTS***

This work was supported by NIH grant AI34155 to D.J.K. We thank Dr. Brenda Peculis at NIH for sharing her buffer recipes and digestion conditions, and Dr. Michael Harris from Case Western University for samples of GMPS and technical advice. We would also like to thank Dr. Laura Yu for helping with experimental procedures and data analyses.

## REFERENCES

- Abraham, J.M., Feagin, J.E., and Stuart, K. 1988. Characterization of cytochrome c oxidase III transcripts that are edited only in the 3' region. *Cell* **55**: 267-272.
- Adler, B.K. and Hajduk, S.L. 1997. Guide RNA requirement for editing-site-specific endonucleolytic cleavage of preedited mRNA by mitochondrial ribonucleoprotein particles in *Trypanosoma brucei*. *Mol. Cell Biol.* **17**: 5377-5385.
- Ammerman, M.L., Fisk, J.C., and Read, L.K. 2008. gRNA/pre-mRNA annealing and RNA chaperone activities of RBP16. *RNA*. **14**: 1069-1080.
- Aphasizhev, R., Aphasizheva, I., Nelson, R.E., and Simpson, L. 2003. A 100-kD complex of two RNA-binding proteins from mitochondria of *Leishmania tarentolae* catalyzes RNA annealing and interacts with several RNA editing components. *RNA* **9**: 62-76.
- Auron, P.E., Weber, L.D., and Rich, A. 1982. Comparison of transfer ribonucleic acid structures using cobra venom and S1 endonucleases. *Biochemistry* **21**: 4700-4706.
- Bernal, J.J., and Garcia-Arenal, F. 1997. Analysis of the in vitro secondary structure of cucumber mosaic virus satellite RNA. *RNA* **3**: 1052-1067.
- Bhat, G.J., Koslowsky, D.J., Feagin, J.E., Smiley, B.L., and Stuart, K. 1990. An extensively edited mitochondrial transcript in kinetoplastids encodes a protein homologous to ATPase subunit 6. *Cell* **61**: 885-894.
- Blom, D., Burg, J., Breek, C.K., Speijer, D., Muijsers, A.O., and Benne, R. 2001. Cloning and characterization of two guide RNA-binding proteins from mitochondria of *Crithidia fasciculata*: gBP27, a novel protein, and gBP29, the orthologue of *Trypanosoma brucei* gBP21. *Nucleic Acids Res* **29**: 2950-2962.
- Blum, B., Bakalara, N., and Simpson, L. 1990. A model for RNA editing in kinetoplastid mitochondria: "guide" RNA molecules transcribed from maxicircle DNA provide the edited information. *Cell* **60**: 189-198.
- Blum, B. and Simpson, L. 1990. Guide RNAs in kinetoplastid mitochondria have a nonencoded 3' oligo(U) tail involved in recognition of the preedited region. *Cell* **62**: 391-397.
- Brunel, C., Romby, P., Westhof, E., Ehresmann, C., and Ehresmann, B. 1991. Three-dimensional model of Escherichia coli ribosomal 5 S RNA as deduced from structure probing in solution and computer modeling. *J. Mol. Biol.* **221**: 293-308.
- Burgess, M.L., Heidmann, S., and Stuart, K. 1999. Kinetoplastid RNA editing does not require the terminal 3' hydroxyl of guide RNA, but modifications to the guide RNA terminus can inhibit in vitro U insertion. *RNA* **5**: 883-892.

- Burgin, A.B. and Pace, N.R. 1990. Mapping the active site of ribonuclease P RNA using a substrate containing a photoaffinity agent. *EMBO J.* **9**: 4111-4118.
- Burkard, M.E., Turner, C.M., and Tinoco, I., Jr. 1999. The interactions that shape RNA structure. In *The RNA world*, 2nd ed. (eds. R.F. Gesteland, T.R. Cech, and J.F. Atkins), pp. 709. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Butcher, S.E., Dieckmann, T., and Feigon, J. 1997. Solution structure of a GAAA tetraloop receptor RNA. *EMBO J.* **16**: 7490-7499.
- Byrne, E.M., Connell, G.J., and Simpson, L. 1996. Guide RNA-directed uridine insertion RNA editing in vitro. *EMBO J.* **15**: 6758-6765.
- Carnes, J., Trotter, J.R., Peltan, A., Fleck, M., and Stuart, K. 2008. RNA editing in *Trypanosoma brucei* requires three different editosomes. *Mol. Cell Biol.* **28**: 122-130.
- Cech, T.R., Tanner, N.K., Tinoco, I., Jr., Weir, B.R., Zuker, M., and Perlman, P.S. 1983. Secondary structure of the *Tetrahymena* ribosomal RNA intervening sequence: structural homology with fungal mitochondrial intervening sequences. *Proc. Natl. Acad. Sci. USA* **80**: 3903-3907.
- Cifuentes-Rojas, C., Halbig, K., Sacharidou, A., De Nova-Ocampo, M., and Cruz-Reyes, J. 2005. Minimal pre-mRNA substrates with natural and converted sites for full-round U insertion and U deletion RNA editing in trypanosomes. *Nucleic Acids Res.* **33**: 6610-6620.
- Cifuentes-Rojas, C., Pavia, P., Hernandez, A., Osterwisch, D., Puerta, C., and Cruz-Reyes, J. 2007. Substrate determinants for RNA editing and editing complex interactions at a site for full-round U insertion. *J. Biol. Chem.* **282**: 4265-4276.
- Connell, G.J. and Simpson, L. 1998. Role of RNA Structure in RNA Editing. In *RNA Structure and Function*. (eds. R.W. Simons, and M. Grunberg-Manago), pp. 641-667. Cold Spring Harbor Laboratory Press, Plainview, NY.
- Corell, R.A., Read, L.K., Riley, G.R., Nellissery, J.K., Allen, T.E., Kable, M.L., Wachal, M.D., Seiwert, S.D., Myler, P.J., and Stuart, K.D. 1996. Complexes from *Trypanosoma brucei* that exhibit deletion editing and other editing-associated properties. *Mol. Cell Biol.* **16**: 1410-1418.
- Cruz-Reyes, J., Rusche, L.N., and Sollner-Webb, B. 1998. *Trypanosoma brucei* U insertion and U deletion activities co-purify with an enzymatic editing complex but are differentially optimized. *Nucleic Acids Res.* **26**: 3634-3639.
- Cruz-Reyes, J., and Sollner-Webb, B. 1996. Trypanosome U-deletional RNA editing involves guide RNA-directed endonuclease cleavage, terminal U exonuclease, and RNA ligase activities. *Proc. Natl. Acad. Sci. USA* **93**: 8901-8906.

- Cruz-Reyes, J., Zhelonkina, A., Rusche, L., and Sollner-Webb, B. 2001. Trypanosome RNA editing: simple guide RNA features enhance U deletion 100-fold. *Mol. Cell Biol.* **21**: 884-892.
- Dubreuil, Y.L., Expert-Bezancon, A., and Favre, A. 1991. Conformation and structural fluctuations of a 218 nucleotides long rRNA fragment: 4-thiouridine as an intrinsic photolabelling probe. *Nucleic Acids Res.* **19**: 3653-3660.
- Ehresmann, C., Baudin, F., Mougel, M., Romby, P., Ebel, J.P., and Ehresmann, B. 1987. Probing the structure of RNAs in solution. *Nucleic Acids Res.* **15**: 9109-9128.
- Favre, A., Saintome, C., Fourrey, J.L., Clivio, P., and Laugaa, P. 1998. Thionucleobases as intrinsic photoaffinity probes of nucleic acid structure and nucleic acid-protein interactions. *J. Photochem. Photobiol. B.* **42**: 109-124.
- Freier, S.M., Kierzek, R., Jaeger, J.A., Sugimoto, N., Caruthers, M.H., Neilson, T., and Turner, D.H. 1986. Improved free-energy parameters for predictions of RNA duplex stability. *Proc. Natl. Acad. Sci. USA* **83**: 9373-9377.
- Golden, D.E. and Hajduk, S.L. 2006. The importance of RNA structure in RNA editing and a potential proofreading mechanism for correct guide RNA:pre-mRNA binary complex formation. *J. Mol. Biol.* **359**: 585-596.
- Harris, M.E. and Christian, E.L. 1999. Use of circular permutation and end modification to position photoaffinity probes for analysis of RNA structure. *Methods* **18**: 51-59.
- Hermann, T., Schmid, B., Heumann, H., and Goring, H.U. 1997. A three-dimensional working model for a guide RNA from *Trypanosoma brucei*. *Nucleic Acids Res.* **25**: 2311-2318.
- Igo, R.P., Jr., Lawson, S.D., and Stuart, K. 2002. RNA sequence and base pairing effects on insertion editing in *Trypanosoma brucei*. *Mol. Cell Biol.* **22**: 1567-1576.
- Igo, R.P., Jr., Palazzo, S.S., Burgess, M.L., Panigrahi, A.K., and Stuart, K. 2000. Uridylate addition and RNA ligation contribute to the specificity of kinetoplastid insertion RNA editing. *Mol. Cell Biol.* **20**: 8447-8457.
- Kable, M.L., Seiwert, S.D., Heidmann, S., and Stuart, K. 1996. RNA editing: a mechanism for gRNA-specified uridylate insertion into precursor mRNA. *Science* **273**: 1189-1195.
- Kang, X., Rogers, K., Gao, G., Falick, A.M., Zhou, S., and Simpson, L. 2005. Reconstitution of uridine-deletion precleaved RNA editing with two recombinant enzymes. *Proc. Natl. Acad. Sci. USA* **102**: 1017-1022.
- Kapushoc, S.T. and Simpson, L. 1999. In vitro uridine insertion RNA editing mediated by cis-acting guide RNAs. *RNA* **5**: 656-669.

- Knapp, G. 1989. Enzymatic approaches to probing of RNA secondary and tertiary structure. *Methods Enzymol.* **180**: 192-212.
- Koller, J., Muller, U.F., Schmid, B., Missel, A., Kruff, V., Stuart, K., and Goring, H.U. 1997. *Trypanosoma brucei* gBP21. An arginine-rich mitochondrial protein that binds to guide RNA with high affinity. *J. Biol. Chem.* **272**: 3749-3757.
- Koslowsky, D.J. 2004. A historical perspective on RNA editing: how the peculiar and bizarre became mainstream. *Methods Mol. Biol.* **265**: 161-197.
- Koslowsky, D.J., Bhat, G.J., Read, L.K., and Stuart, K. 1991. Cycles of progressive realignment of gRNA with mRNA in RNA editing. *Cell* **67**: 537-546.
- Koslowsky, D.J., Reifur, L., Yu, L.E., and Chen, W. 2004. Evidence for U-tail stabilization of gRNA/mRNA interactions in kinetoplastid RNA editing. *RNA Biol.* **1**: 28-34.
- Krol, A., Westhof, E., Bach, M., Luhrmann, R., Ebel, J.P., and Carbon, P. 1990. Solution structure of human U1 snRNA. Derivation of a possible three-dimensional model. *Nucleic Acids Res.* **18**: 3803-3811.
- Lawson, S.D., Igo, R.P., Jr., Salavati, R., and Stuart, K.D. 2001. The specificity of nucleotide removal during RNA editing in *Trypanosoma brucei*. *RNA* **7**: 1793-1802.
- Lescoute, A. and Westhof, E. 2006. Topology of three-way junctions in folded RNAs. *RNA* **12**: 83-93.
- Leung, S.S. and Koslowsky, D.J. 1999. Mapping contacts between gRNA and mRNA in trypanosome RNA editing. *Nucleic Acids Res.* **27**: 778-787.
- Leung, S.S. and Koslowsky, D.J. 2001a. Interactions of mRNAs and gRNAs involved in trypanosome mitochondrial RNA editing: structure probing of an mRNA bound to its cognate gRNA. *RNA* **7**: 1803-1816.
- Leung, S.S. and Koslowsky, D.J. 2001b. RNA editing in *Trypanosoma brucei*: characterization of gRNA U-tail interactions with partially edited mRNA substrates. *Nucleic Acids Res.* **29**: 703-709.
- Lockard, R.E. and Kumar, A. 1981. Mapping tRNA structure in solution using double-strand-specific ribonuclease V1 from cobra venom. *Nucleic Acids Res.* **9**: 5125-5140.
- Lowman, H.B. and Draper, D.E. 1986. On the recognition of helical RNA by cobra venom V1 nuclease. *J. Biol. Chem.* **261**: 5396-5403.
- Madison-Antenucci, S., Grams, J., and Hajduk, S.L. 2002. Editing machines: the complexities of trypanosome RNA editing. *Cell* **108**: 435-438.

- Maslov, D.A. and Simpson, L. 1992. The polarity of editing within a multiple gRNA-mediated domain is due to formation of anchors for upstream gRNAs by downstream editing. *Cell* **70**: 459-467.
- Mathews, D.H., Disney, M.D., Childs, J.L., Schroeder, S.J., Zuker, M., and Turner, D.H. 2004. Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. *Proc. Natl. Acad. Sci. USA* **101**: 7287-7292.
- Miller, M.M., Halbig, K., Cruz-Reyes, J., and Read, L.K. 2006. RBP16 stimulates trypanosome RNA editing in vitro at an early step in the editing reaction. *RNA* **12**: 1292-1303.
- Milligan, J.F., Groebe, D.R., Witherell, G.W., and Uhlenbeck, O.C. 1987. Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res.* **15**: 8783-8798.
- Milligan, J.F. and Uhlenbeck, O.C. 1989. Synthesis of small RNAs using T7 RNA polymerase. *Methods Enzymol.* **180**: 51-62.
- Muller, U.F. and Goring, H.U. 2002. Mechanism of the gBP21-mediated RNA/RNA annealing reaction: matchmaking and charge reduction. *Nucleic Acids Res.* **30**: 447-455.
- Panigrahi, A.K., Allen, T.E., Stuart, K., Haynes, P.A., and Gygi, S.P. 2003. Mass spectrometric analysis of the editosome and other multiprotein complexes in *Trypanosoma brucei*. *J Am Soc Mass Spectrom* **14**: 728-735.
- Panigrahi, A.K., Ernst, N.L., Domingo, G.J., Fleck, M., Salavati, R., and Stuart, K.D. 2006. Compositionally and functionally distinct editosomes in *Trypanosoma brucei*. *RNA* **12**: 1038-1049.
- Panigrahi, A.K., Schnauffer, A., and Stuart, K.D. 2007. Isolation and compositional analysis of trypanosomatid editosomes. *Methods Enzymol.* **424**: 3-24.
- Peattie, D.A. 1979. Direct chemical method for sequencing RNA. *Proc. Natl. Acad. Sci. USA* **76**: 1760-1764.
- Peterson, R.D., and Feigon, J. 1996. Structural change in Rev responsive element RNA of HIV-1 on binding Rev peptide. *J. Mol. Biol.* **264**: 863-877.
- Pollard, V.W., Harris, M.E., and Hajduk, S.L. 1992. Native mRNA editing complexes from *Trypanosoma brucei* mitochondria. *EMBO J.* **11**: 4429-4438.
- Reenan, R.A. 2005. Molecular determinants and guided evolution of species-specific RNA editing. *Nature* **434**: 409-413.

- Rusche, L.N., Cruz-Reyes, J., Piller, K.J., and Sollner-Webb, B. 1997. Purification of a functional enzymatic editing complex from *Trypanosoma brucei* mitochondria. *EMBO J.* **16**: 4069-4081.
- Sacharidou, A., Cifuentes-Rojas, C., Halbig, K., Hernandez, A., Dangott, L.J., De Nova-Ocampo, M., and Cruz-Reyes, J. 2006. RNA editing complex interactions with a site for full-round U deletion in *Trypanosoma brucei*. *RNA* **12**: 1219-1228.
- Schmid, B., Riley, G.R., Stuart, K., and Goring, H.U. 1995. The secondary structure of guide RNA molecules from *Trypanosoma brucei*. *Nucleic Acids Res.* **23**: 3093-3102.
- Schnauffer, A., Ernst, N.L., Palazzo, S.S., O'Rear, J., Salavati, R., and Stuart, K. 2003. Separate insertion and deletion subcomplexes of the *Trypanosoma brucei* RNA editing complex. *Mol. Cell* **12**: 307-319.
- Schroeder, R., Barta, A., and Semrad, K. 2004. Strategies for RNA folding and assembly. *Nat Rev Mol. Cell Biol.* **5**: 908-919.
- Schumacher, M.A., Karamooz, E., Zikova, A., Trantirek, L., and Lukes, J. 2006. Crystal structures of *T. brucei* MRP1/MRP2 guide-RNA binding complex reveal RNA matchmaking mechanism. *Cell* **126**: 701-711.
- Seiwert, S.D., Heidmann, S., and Stuart, K. 1996. Direct visualization of uridylyate deletion in vitro suggests a mechanism for kinetoplastid RNA editing. *Cell* **84**: 831-841.
- Seiwert, S.D. and Stuart, K. 1994. RNA editing: transfer of genetic information from gRNA to precursor mRNA in vitro. *Science* **266**: 114-117.
- Simpson, L., Aphasizhev, R., Gao, G., and Kang, X. 2004. Mitochondrial proteins and complexes in *Leishmania* and *Trypanosoma* involved in U-insertion/deletion RNA editing. *RNA* **10**: 159-170.
- Simpson, L., Sbicego, S., and Aphasizhev, R. 2003. Uridine insertion/deletion RNA editing in trypanosome mitochondria: a complex business. *RNA* **9**: 265-276.
- Sontheimer, E.J. 1994. Site-specific RNA crosslinking with 4-thiouridine. *Mol. Biol. Rep.* **20**: 35-44.
- Stuart, K.D., Schnauffer, A., Ernst, N.L., and Panigrahi, A.K. 2005. Complex management: RNA editing in trypanosomes. *Trends Biochem. Sci.* **30**: 97-105.
- Sturm, N.R. and Simpson, L. 1990. Partially edited mRNAs for cytochrome b and subunit III of cytochrome oxidase from *Leishmania tarentolae* mitochondria: RNA editing intermediates. *Cell* **61**: 871-878.
- Thomas, B.C., Kazantsev, A.V., Chen, J.L., and Pace, N.R. 2000. Photoaffinity cross-linking and RNA structure analysis. *Methods Enzymol.* **318**: 136-147.

- Vondruskova, E., van den Burg, J., Zikova, A., Ernst, N.L., Stuart, K., Benne, R., and Lukes, J. 2005. RNA interference analyses suggest a transcript-specific regulatory role for mitochondrial RNA-binding proteins MRP1 and MRP2 in RNA editing and other RNA processing in *Trypanosoma brucei*. *J. Biol. Chem.* **280**: 2429-2438.
- Walter, A.E., Turner, D.H., Kim, J., Lyttle, M.H., Muller, P., Mathews, D.H., and Zuker, M. 1994. Coaxial stacking of helices enhances binding of oligoribonucleotides and improves predictions of RNA folding. *Proc. Natl. Acad. Sci. USA* **91**: 9218-9222.
- Westhof, E., Romby, P., Romaniuk, P.J., Ebel, J.P., Ehresmann, C., and Ehresmann, B. 1989. Computer modeling from solution data of spinach chloroplast and of *Xenopus laevis* somatic and oocyte 5 S rRNAs. *J. Mol. Biol.* **207**: 417-431.
- Yu, L.E. and Koslowsky, D.J. 2006. Interactions of mRNAs and gRNAs involved in trypanosome mitochondrial RNA editing: structure probing of a gRNA bound to its cognate mRNA. *RNA* **12**: 1050-1060.
- Zikova, A., Kopečna, J., Schumacher, M.A., Stuart, K., Trantirek, L., and Lukes, J. 2008. Structure and function of the native and recombinant mitochondrial MRP1/MRP2 complex from *Trypanosoma brucei*. *Int. J. Parasitol.* **31**: 901-912.
- Zimmermann, G.R., Jenison, R.D., Wick, C.L., Simorre, J.P., and Pardi, A. 1997. Interlocking structural motifs mediate molecular discrimination by a theophylline-binding RNA. *Nat. Struct. Biol.* **4**: 644-649.
- Znosko, B.M., Kennedy, S.D., Wille, P.C., Krugh, T.R., and Turner, D.H. 2004. Structural features and thermodynamics of the J4/5 loop from the *Candida albicans* and *Candida dubliniensis* group I introns. *Biochemistry* **43**: 15822-15837.
- Zuker, M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**: 3406-3415.

## **Chapter 3**

# **THE IMPACT OF MRNA STRUCTURE ON GUIDE RNA TARGETING IN KINETOPLASTID RNA EDITING**

Larissa Reifur, Laura E. Yu, Jorge Cruz-Reyes, Michelle vanHartesvelt and Donna J. Koslowsky 2008. Submitted to RNA Journal.

## **ABSTRACT**

Mitochondrial mRNA editing in *Trypanosoma brucei* requires the specific interaction of a guide RNA with its cognate mRNA. Hundreds of different gRNAs are involved in the editing process, each needing to target their specific editing domain within their target message. We hypothesized that the structure surrounding the mRNA target may be a limiting factor and involved in the regulation process. In this study, we selected four mRNAs with distinct predicted secondary structures and investigated how sequence and structure of both the gRNA and mRNA affected efficient gRNA targeting. Two of the mRNAs, including the A6 and ND7-550 (5' end of NADH dehydrogenase subunit 7) that have open, accessible anchor binding sites, show very efficient gRNA targeting. Through electrophoretic mobility shift assays, we found that the cognate gRNA for ND7-550 had higher affinity for its mRNA than the gRNA for the A6 mRNA, which was explained by its faster association rate constant determined by surface plasmon resonance studies. In contrast, mRNAs with considerable structure surrounding the anchor binding sites were less accessible and had very low affinity for their cognate gRNAs. For these substrate pairs, effective interaction would require protein cofactors, introducing a step that would allow for regulation.

## **INTRODUCTION**

Target RNA structure can affect the access of non-coding complementary oligonucleotides and consequently, regulate gene expression in a wide range of organisms (Lima et al. 1992; Campbell et al. 1997; Eckardt et al. 1997; Franch et al. 1999; Patzel et al. 1999; Amarzguioui et al. 2000; Vickers et al. 2000; Mercatanti et al. 2002; Walton et al. 2002; Kikuchi et al. 2003; Brown et al. 2005; Hackermuller et al. 2005; Overhoff et al.

2005; Muckstein et al. 2006; Kertesz et al. 2007; Long et al. 2008; Tafer et al. 2008). In prokaryotes, natural antisense RNAs are involved in regulating plasmid copy number, transposons and bacteriophages (Wagner and Simons 1994; Wassarman et al. 1999; Argaman et al. 2001; Ying et al. 2008). In eukaryotes, large numbers of miRNAs have been identified that can target specific mRNAs, inducing degradation or translation suppression (Brennecke et al. 2005; Lewis et al. 2005; Ying et al. 2008). With the onset of RNAi technologies, it was found that target sequence and structure were fundamental for specific and efficient interaction of RNA ligands (Long et al. 2007; Shao et al. 2007a; Tafer et al. 2008). We are interested in mitochondrial mRNA editing in *Trypanosoma brucei*, another pathway for gene regulation that similarly involves mRNA targeting and RNA-RNA interaction.

Trypanosome mitochondrial RNA editing involves insertion and deletion of uridylates (Us) to generate functional open reading frames (Benne et al. 1986; Madison-Antenucci et al. 2002; Simpson et al. 2003; Stuart et al. 2005). Editing is catalyzed by a multiprotein complex, the editosome (Carnes et al. 2008) and requires small guide RNAs (gRNAs) to precisely direct the events (Seiwert and Stuart 1994). The amount of RNA editing required is substantial, involving 12 genes and at least 1200 gRNAs (Benne 1994; Stuart et al. 2005; Ochsenreiter et al. 2008). Guide RNAs are key components of the editing reaction, as they interact with the mRNA, supply the information for the nucleotide alterations and appear to be able to direct the cleavage and ligation events (Seiwert and Stuart 1994). They have an average length of 50–70 nt with approximately 15 Us added post-transcriptionally at the 3' end (Blum and Simpson 1990). Their 5' end, the gRNA anchor (4 to 16 nt), is responsible for selection and binding to the pre-edited

mRNA (Blum and Simpson 1992). The target mRNA contains an anchor binding site (ABS), with sequence complementary to the gRNA anchor, located just 3' to the editing domain (Blum et al. 1990). Each gRNA must hybridize to its complementary “anchor” sequence in the mRNA target. While structural studies of the gRNAs indicate that they fold into a common structure (Schmid et al. 1995; Hermann et al. 1997), the hundreds of different targeting sites within the mRNAs suggest that the structures surrounding the anchor binding sites will affect the efficiency and speed of the targeting event.

In previous work using electrophoretic mobility shift assays (EMSA), we were able to show that the secondary structures of the interacting RNAs play significant roles in complex affinity and stability (Koslowsky et al. 2004). The unpaired CYb mRNA forms a stable stem-loop that includes the ABS in a mostly double-stranded stem. The observed binding affinity for this substrate paired with its cognate gRNA, gCYb-558, was very low ( $K_D = 0.6 \mu\text{M}$ ). In contrast, the 3' end of A6 is much less structured and its ABS spans a terminal loop and one side of a short (7 bp) stem. The observed binding affinity for the A6/gA6-14 pair was much higher ( $K_D = 5.8 \text{ nM}$ ) than for the CYb pair. We suggest that the stable structural features of CYb, especially the double-stranded nature of its ABS, is likely responsible for the poor gRNA interaction. We hypothesize that structural features surrounding the ABS in the mRNA may limit gRNA access and play an important role in the developmental regulation of RNA editing. In the present work, we utilized four mRNA/gRNA pairs to investigate how structural features surrounding the ABS region may influence gRNA binding and the RNA editing process.

## **RESULTS**

### **Description of mRNAs and analyses of their predicted secondary structures**

The four mRNA/gRNA pairs were used in this study included A6U/gA6-14, CYbU/gCYb-558, ND7UHR3/gND7-506, and ND7-550/gND7-550 (Fig. 3.1). Both the A6 and CYb pairs have been previously described (Koslowsky et al. 1996). ND7UHR3 and ND7-550 are from two separate regions within the 5' editing domain of the ND7 mRNA and require different gRNAs to be edited (Koslowsky et al. 1990). ND7UHR3 is located just 5' of the homology region three (HR3) of the ND7 mRNA. ND7HR3 contains the anchor sequence for the 5' domain initiating gRNA, gND7-506. It is 98 nt long, including 46 nt upstream and 36 nt downstream of the ABS. ND7-550 is a 76 nt-long substrate located at the 5'-end of ND7 and contains the ABS for gND7-550. This ABS is a span of 14 nucleotides that are not edited in the mature transcript. The gND7-550 gRNA was originally cloned and sequenced as part of a gRNA/mRNA chimeric molecule that had been isolated from a pool of partially edited ND7 molecules (Koslowsky et al. 1991). The isolated cDNA was partially edited as directed by gND7-550, but unedited 3' to the anchor-binding site, suggesting that gND7-550 could initiate the editing process. Like A6, the two selected regions of the ND7 mRNA are constitutively and extensively edited. While gCYb-558 can also initiate a cascade of editing events, the CYb mRNA is not edited in the slender bloodstream stage of the trypanosome life cycle (Feagin et al. 1987).

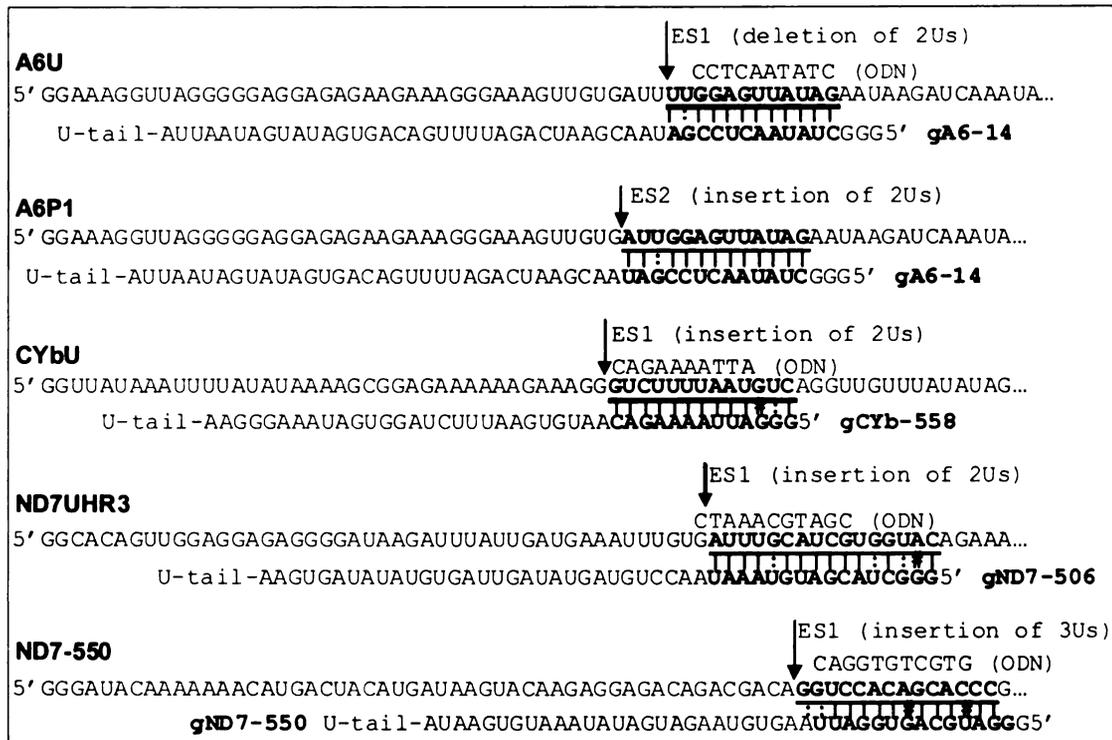


Figure 3.1 mRNAs aligned with gRNAs and ODNs.

The mRNA anchor binding site (**bold and underlined**) is complementary to the gRNA anchor and ODN. Watson-Crick (|), non-Watson-Crick (:), mismatches (#) and the first editing site (ES) are indicated. The mRNA sequences continue at the 3'-end, as in figure 3.2.

Computer modeling of the mRNAs indicate that they have distinct secondary structures around the ABS (Fig. 3.2). A6U forms the least stable structure and presents most of the anchor binding site within a terminal loop, defined by an 8 bp stem ( $\Delta G_{27^\circ\text{C}} = -8 \text{ Kcal mole}^{-1}$ ). In contrast, CYbU forms a stable stem-loop with its ABS located mostly in a double-stranded region within the stem ( $\Delta G_{27^\circ\text{C}} = -24.5 \text{ Kcal mole}^{-1}$ ). The structures obtained by enzymatic and chemical solution structure probing of A6U (Reifur and Koslowsky 2008) and CYbU (Leung and Koslowsky 2001a) support the computer predicted structures. Both ND7 mRNA substrates had predicted structures that were less stable than CYbU (ND7UHR3,  $\Delta G_{27^\circ\text{C}} = -15.9 \text{ kcal mole}^{-1}$  and ND7-550,  $\Delta G_{27^\circ\text{C}} = -10.8$

kcal mole<sup>-1</sup>). ND7UHR3 was predicted to have the ABS within a double-stranded region with internal loops, while the ABS for ND7-550 was mainly in a single-stranded region.

According to these structure models, ND7-550, but not ND7UHR3 or CYbU, is predicted to have ABS accessibility and binding affinity similar to A6U.

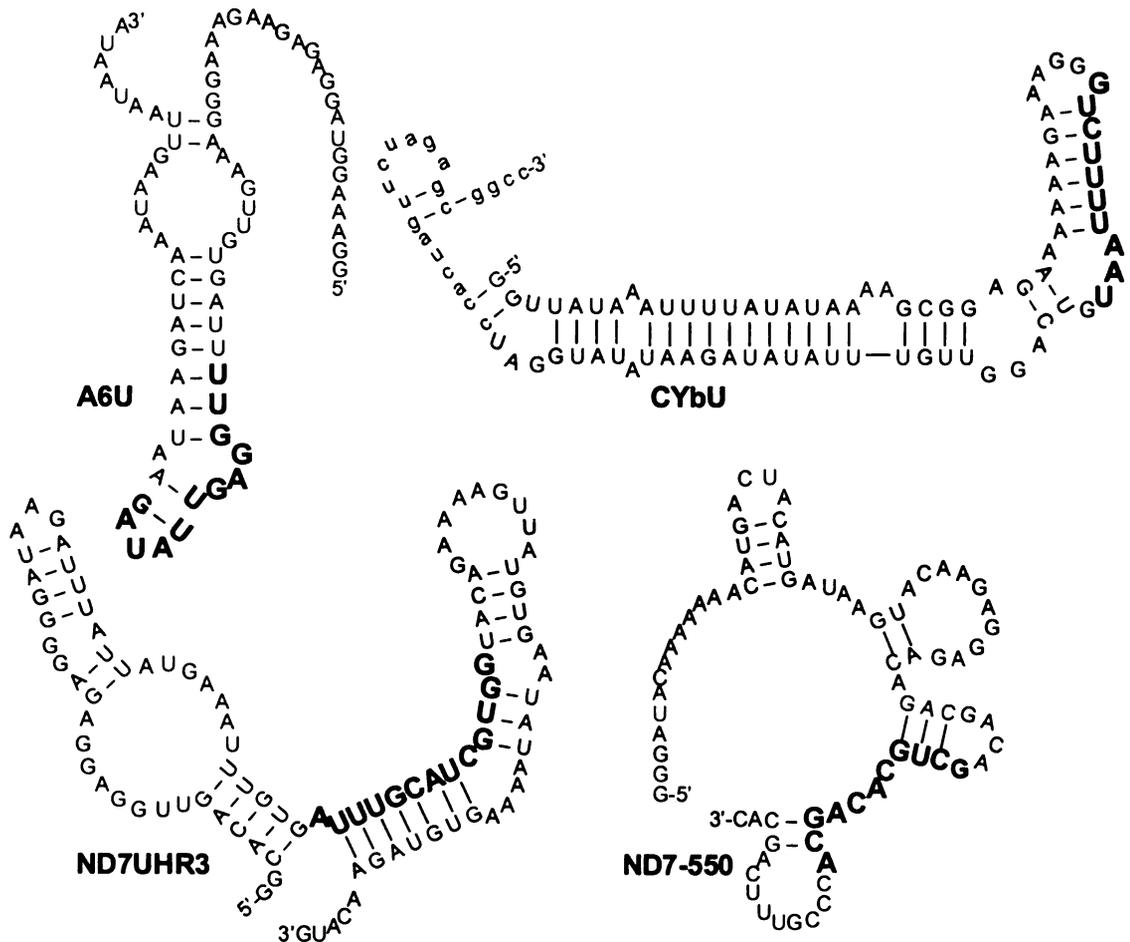


Figure 3.2 Predicted secondary structures for A6U, CYbU, ND7UHR3, ND7-550. The anchor binding site (ABS) is in bold.

### Determination of ABS accessibility

To indirectly measure ABS accessibility, we conducted RNase H-mediated cleavage assays as previously described (Birikh et al. 1997; Amarzguioui et al. 2000). In these experiments, 10 pmols of 5'-<sup>32</sup>P-labeled mRNA were renatured in vitro (40 mM Tris-HCl

pH 7.5, 100 mM KCl, 2 mM MgCl<sub>2</sub>, 1 mM DTT). RNase H and a 10-11 nt oligodeoxyribonucleotide (ODN) complementary to the ABS were concomitantly added to the folded mRNA and aliquots were taken for analysis. The relative ABS accessibility was evaluated by using different ratios of target mRNA to ODN, including 1:1, 1:10, and 1:30 ratios, and by assaying for cleavage at three different time points (1, 15, and 30 minutes). If the ODN was able to bind the ABS, RNase H cleaved within the DNA-RNA duplex. The reactions were specific and demonstrated reproducible and expected cleavage products (Fig. 3.3, left column). Under the conditions utilized, quantitative analyses of substrate cleavage revealed different degrees of digestion, depending on the mRNA, the incubation time, and the ODN concentration (Fig. 3.3, data not shown). As predicted, the two substrates with ABS within a single-stranded region (A6U and ND7-550) were the most accessible, showing the highest ODN-directed cleavage. Very interestingly, A6U was the only substrate cleaved at the shortest time point (1 min). However, as the time of incubation increased, the percentage of digested ND7-550 surpassed the maximum percentage of A6 cleavage by approximately 10% (Fig. 3.3B). In contrast, the CYbU substrate showed no cleavage at even the highest ODN concentration, indicating the energetic difficulty involved in invading the stable mRNA stem-loop structure. Cleavage of the ND7UHR3 substrate was observed at the lowest (1:1) ratios and the percentage of digested ND7UHR3 did increase with increasing amounts of ODN and time of digestion. Nevertheless, its maximum digestion was substantially lower than the percentage of A6U cleaved at the lower mRNA to ODN ratios (Fig. 3.3B). These results suggest that this assay could be used to quickly assess the accessibility of specific gRNA targets.

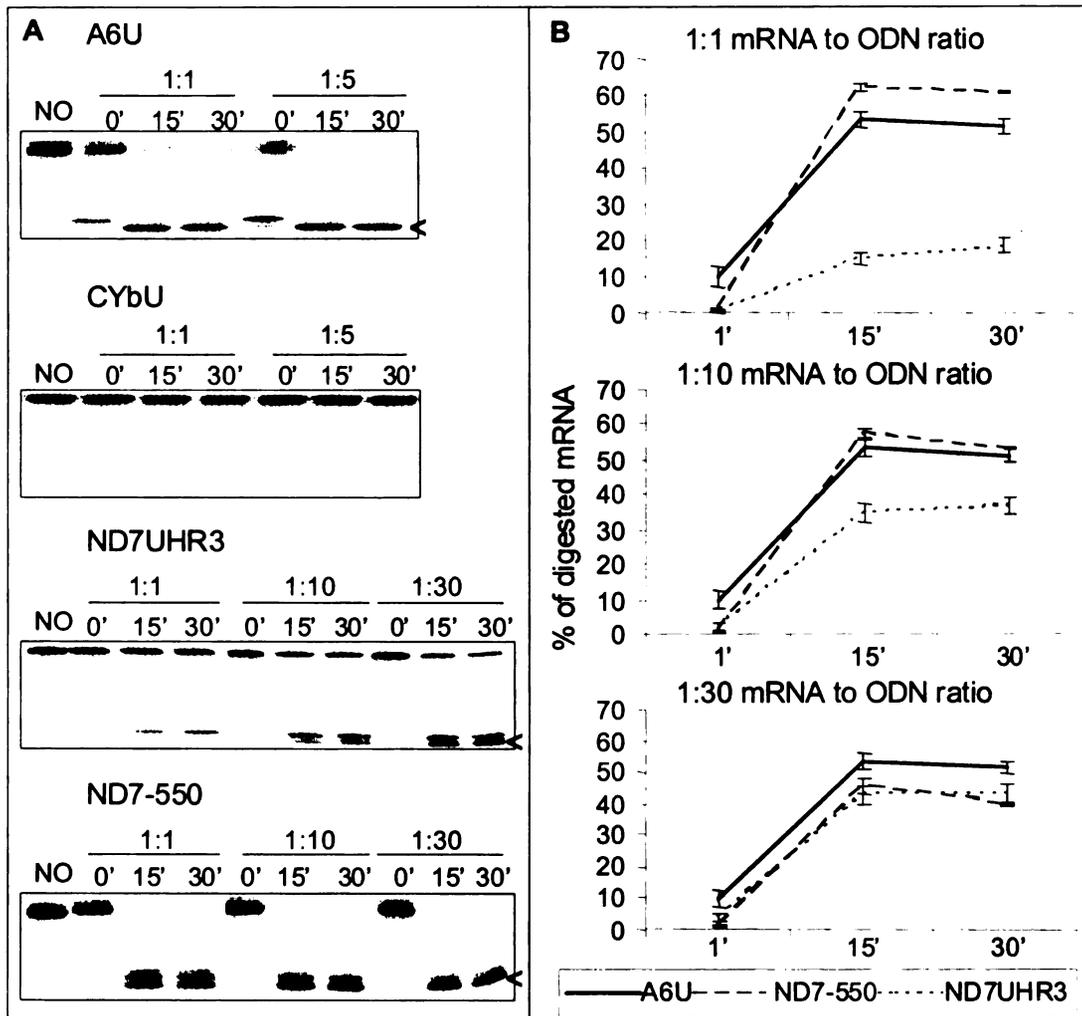


Figure 3.3 ODN-directed accessibility assays.

A. representative images of 8% denaturing polyacrylamide gels. Each reaction contained a pre-hybridized  $^{32}\text{P}$ -labeled mRNA (A6U, CYbU, ND7UHR3, or ND7-550) that was digested with RNase H for 1, 15, and 30 minutes upon addition of a specific ODN (1:1, 1:5, 1:10, or 1:30 mRNA to ODN ratio). "NO": no ODN control. The digested products (<) are indicated. B. percentage of RNase H digestion products. The amount of digested A6U shown in each graph was kept constant at 1:1 ratio. The CYbU mRNA was not included because no digested products were detected. These data are the average of three experiments.

### Analysis of binding affinities

To correlate ABS accessibility with strength of complex formation we determined the binding affinity for the mRNA/gRNA pairs using EMSA, as previously described

(Koslowsky et al. 2004). In these experiments, 5 nM of  $^{32}\text{P}$ -labeled gRNA, either with or

without its U-tail, was annealed with increasing concentrations of mRNA in a buffer containing 2 mM Mg<sup>++</sup>. The cognate RNAs were combined, denatured at 70 °C for 2 min, slow cooled to 27 °C and allowed to anneal for 3 hours. The free RNAs were then separated from the bound complex by electrophoresis on nondenaturing 6% polyacrylamide gels. For all gRNA/mRNA pairs a single predominant band was observed (Fig. 3.4). Complex formation was quantified on Molecular Dynamics phosphorimager and the apparent dissociation equilibrium constant ( $K_D$ ) calculated. Surprisingly, the apparent  $K_D$  for the ND7-550 pair was almost 10 fold lower than the  $K_D$  measured for the A6 pair ( $K_{D-ND7-550} = 0.3 \pm 0.2$  nM vs  $K_{D-A6U} = 2.7 \pm 0.5$  nM). As predicted, the binding affinity for the ND7UHR3 was weaker,  $K_D = 84.5 \pm 7.6$  nM, but still considerably better than CYb (Table 3.1). In addition, for all gRNA/mRNA pairs, the U-tail increased the affinity of the gRNA for its cognate mRNA. The U-tail contribution was minimal for the A6U/gA6-14 interaction, decreasing the observed  $K_D$  in approximately 2-fold. In contrast, a difference in binding affinity of over 10-fold was observed in the presence or absence of the U-tail for the ND7-550 pair. This indicates that the gRNA U-tail can contribute to the binding affinity even for those gRNAs that show high affinity for their targets.

Table 3.1 mRNA/gRNA apparent binding affinities. The  $K_D$  was also calculated when the gRNAs had the U-tail deleted (No U-tail).

mRNA/gRNA	$K_D$	$K_D$ No U-tail	$K_{rel}$ ( $K_D$ No U-tail/ $K_D$ )
ND7-550/gND7-550	$0.3 \pm 0.2$ nM	$3.9 \pm 0.4$ nM	13
A6U/gA6-14	$2.7 \pm 0.5$ nM	$5.8 \pm 0.4$ nM	2.15
ND7UHR3/gND7-506	$63.8 \pm 6.7$ nM	$175.2 \pm 9.7$ nM	2.74
CYbU/gCYb-558	$0.6 \pm 0.1$ $\mu$ M	$1.3 \pm 0.1$ $\mu$ M	2.16

\* All pairs were hybridized 3 h at 27°C, under identical conditions.

Figure 3.4. mRNA/gRNA binding affinity by EMSA. **A-D**: Representative images of 6% polyacrylamide gels and corresponding binding isotherms (see Table 3.1 for dissociation constants). Samples contained 5 nM  $^{32}$ P-labeled gRNA and increasing concentrations of the cognate mRNA. For the ND7-550 (**A**) and CYbU (**D**) gels, the odd number lanes are for gRNA containing its U-tail ( $\circ$ ) and even number lanes for the gRNA with deleted U-tail ( $\square$ ). For the A6U (**B**) and ND7UHR3 (**C**), the gRNA +U-tail and with no U-tail were separated. Complex formation ( $\blacklozenge$ ) was quantitated as in Koslowsky et al. (2004). The binding isotherms show the average result obtained from 4 experiments for each pair of mRNA/gRNA. Error bars indicate the standard deviation in complex formation.

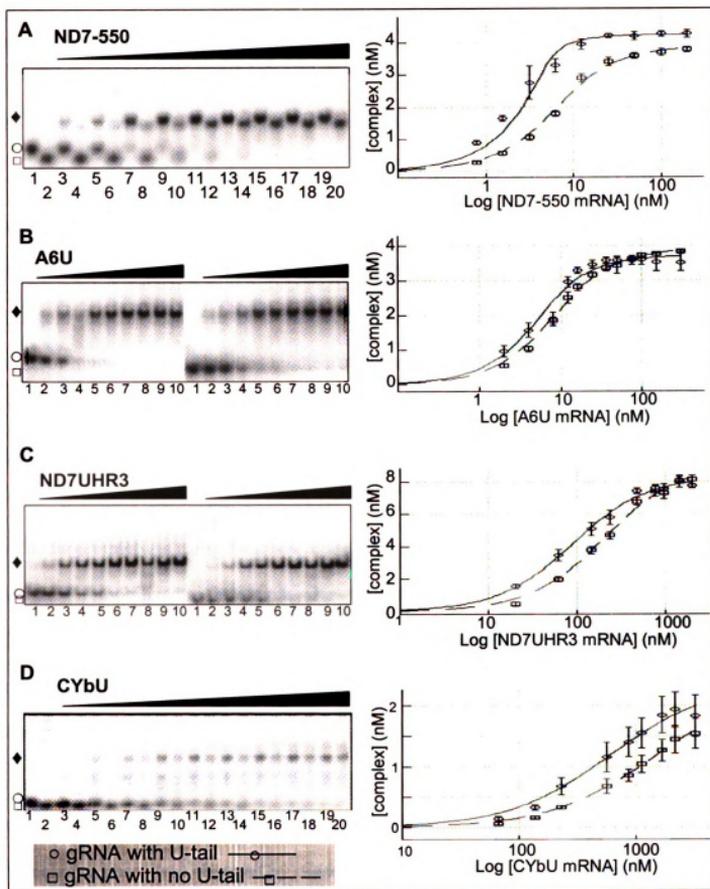


Figure 3.4 mRNA/gRNA binding affinity by EMSA.

## **mRNA/gRNA rate constants**

The efficiency of RNA-dependent systems has been correlated to fast annealing kinetics (Patzel and Sczakiel 1998). To define the association ( $k_{on}$ ) and dissociation ( $k_{off}$ ) rate constants of the most accessible mRNAs, surface plasmon resonance (SPR) was employed. The BIACORE 2000 instrument (Biacore, Uppsala, Sweden) monitors intermolecular interactions through SPR that arises when incident light is reflected from a thin gold film. A change in the refractive index due to molecular binding changes the angle of light, which is recorded by the detector as a change in resonance units (RU). The rate at which this change occurs provides information about the association and dissociation rates of the molecules (Katsamba et al. 2002).

In these experiments, an ODN-tag with a 3' biotin label was ligated to the 3'-ends of the target mRNAs using T4 DNA ligase and a bridge ODN. The biotin labeled mRNA was then immobilized to the streptavidin covered surface of the SA chip. To see reliable gRNA binding to the mRNA, 50 to 350 RU of mRNA were attached to the chip surface, in two of the four channels. One channel remained empty to be used as a reference and one contained the biotinylated tag as control for background binding. A continuous flow of gRNA in binding buffer (100 mM Tris pH 7.5, 0.1 mM EDTA, 2 mM  $MgCl_2$ , and 100 mM KCl) was injected over the immobilized mRNAs to monitor association. The dissociation phase was obtained by chasing the gRNA with buffer. The long injection times and the regeneration procedures used between two binding assays progressively affect the mRNA integrity during the experiments (von der Haar and McCarthy 2003). These limitations made it difficult to generate curves amenable to simple Scatchard-type analyses. Because line fitting using global analysis requires extremely high quality data,

this method of analysis proved impractical (Myszka 1997). Using the separate fit function of Biaevaluation 3.0 (Biacore, Uppsala, Sweden) with a global analysis for each association and dissociation curve separately, allowed an analysis of the individual rate constants. The individual rate constants were averaged from a minimum of three separate experiments (three mRNA concentrations per experiment), and the equilibrium dissociation constant was calculated from the rate constants. The errors reported are based on the variances of all curves obtained (Nordgren et al. 2001). Figure 3.5 shows representative binding curves for the A6U and the ND7-550 interactions. For both of these pairs, the SPR analyses indicate that the gRNA/mRNA interactions are stable with a slow dissociation rate of  $\sim 3.0 \times 10^{-5} \text{ s}^{-1}$ . In contrast, the association rates did differ significantly, with the gND7-550 gRNA binding its target almost 5 fold faster than gA6-14 ( $5.1 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$  vs  $1.2 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$ ). Using the measured rate constants, the affinity constants ( $K_D$ ) for both RNA pairs were calculated to be 2.5 nM and 0.56 nM for the A6U and ND7-550 pairs, respectively. The calculated  $K_D$ s were similar to the  $K_D$ s observed by EMSA, indicating that the increase in affinity observed for the ND7 pair was due to the difference in association rate.

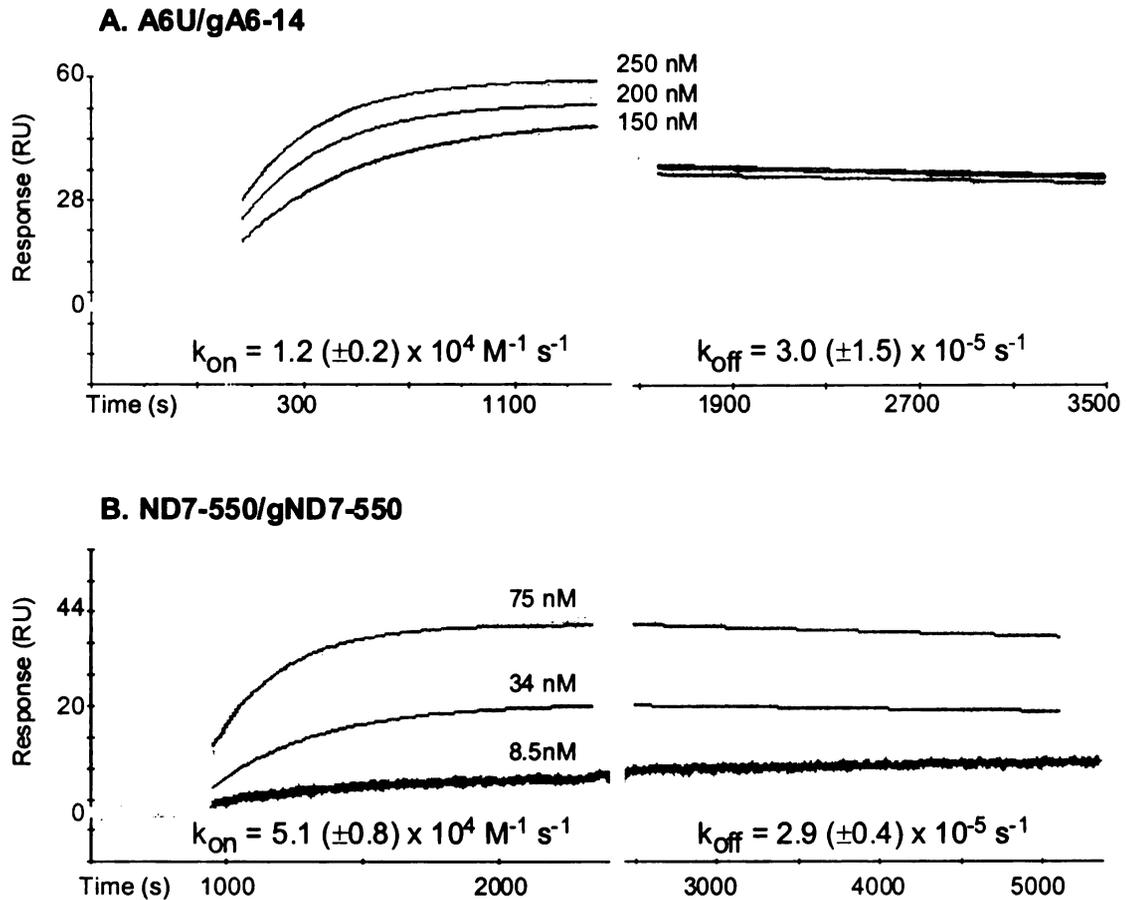


Figure 3.5 mRNA/gRNA sensorgrams and rate constants by Surface Plasmon Resonance. Representative SPR sensorgrams are shown with line fits. A. Sensorgram of A6U + gA6-14. B. Sensorgram of ND7-550 + gND7-550. The association ( $k_{on}$ ) and dissociation ( $k_{off}$ ) rate constants represent the mean of a minimum of 3 runs (each run utilizing 3 different mRNA concentrations) and are listed with the error in parentheses. RU = resonance units.

### Ability of each mRNA/gRNA pair to be recognized and cleaved by the editosome

The 3'-end of the A6 *T. brucei* mRNA is used as a standard in editing reactions in vitro because it is more efficiently edited and undergoes a full round without the need of sequence alteration (Kable et al. 1996; Seiwert et al. 1996; Cruz-Reyes et al. 2001; Lawson et al. 2001; Cifuentes-Rojas et al. 2007; Hernandez et al. 2008). Using standard editing reaction conditions, we evaluated the ability of the mRNA/gRNA pairs used in

this study to undergo cleavage in a gRNA-directed cleavage reaction. The assay consisted of incubating the 5' -<sup>32</sup>P-labeled mRNAs with the respective cognate gRNAs in the presence of mitochondrial extract containing the editing machinery. To obtain the extract, mitochondria from a *T. brucei* culture were isolated, lysed, and fractionated in a glycerol gradient or purified on Q-Sepharose chromatographic column. To standardize the reactions to be of the insertion type, the A6 mRNA was partially edited at its first site (A6P1, Fig. 3.1). To increase cleavage efficiency, free UTP was not added, and the purified mitochondrial extract was pre-treated with inorganic pyrophosphate. All four mRNA/gRNA pairs, in a molar ratio of 1:10, were pre-hybridized and incubated with the above mitochondrial extracts for 1 h, at 27°C. Under these conditions, gRNA-dependent cleavages at editing sites were observed only in A6P1 and ND7-550 (Fig. 3.6). As expected, A6P1 was cleaved at the first expected site (ES2) in a gRNA-directed manner. A second cleavage product was observed 5' of ES2, however it was also present in the absence of the gRNA indicating the presence of a “nonspecific” endonuclease. The cleavages within ND7-550 were also gRNA dependent (note absence of cleavage product when exogenous gRNA was not added). However, five different cleavage sites, C1-5, were observed, none of them at the expected first editing site ES1 (Figs. 3.6, 3.7). C1, 2, and 3 occurred with varied efficiencies between independent assays (data not shown). The C4 and C5 cleavages however, were reproducible and almost as efficient as the cleavage observed at ES2 in the A6 mRNA [efficiencies of approximately 3% (C4) and 1.8% (C5) in ND7-550 and 4% in the A6].

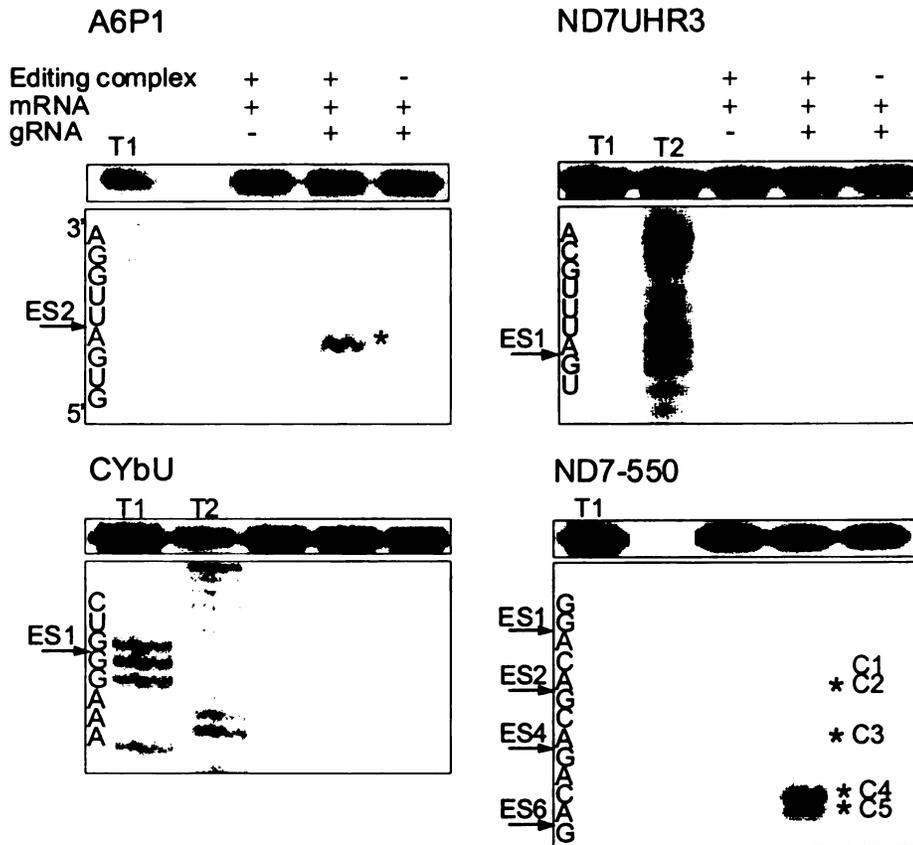


Figure 3.6 In vitro gRNA-directed cleavage assay. Representative images of 8% denaturing polyacrylamide gels. Radiolabeled mRNAs were incubated with their cognate gRNAs (gA6-14, gCYb-558, gND7-506, and gND7-550) in standard cleavage conditions with purified editosomes. (\*) gRNA-directed cleavage products. The cleavage products migrate  $\frac{1}{2}$  nucleotide slower than the products generated with RNase T1 digestion (Cruz-Reyes and Sollner-Webb 1996) due to differences in enzyme site of cleavage. (ES) editing sites, indicated by arrows. (T1 and T2) RNase T1 and T2 digests for ES mapping.

Experiments using the enriched glycerol fractions or the purified editosomes gave identical results (data not shown). Computer modeling of ND7-550/gND7-550 indicates that the cleavages occurred within a single-stranded region that is punctuated with three cytosine residues located just upstream of the anchor duplex. The U-tail is predicted to base pair farther upstream with an uninterrupted run of 9 purines (Fig. 3.7). To insure that the observed cleavages were not due to non-specific single strand RNases present in the mitochondrial preparations, gND7-550 was annealed to ND7-550 in NE buffer 2 at 27°C and the complex's structure was probed with Mung Bean Nuclease (MBN). No digestion by the single stranded specific MBN was observed within the C1-5 region (Fig. 3.8).



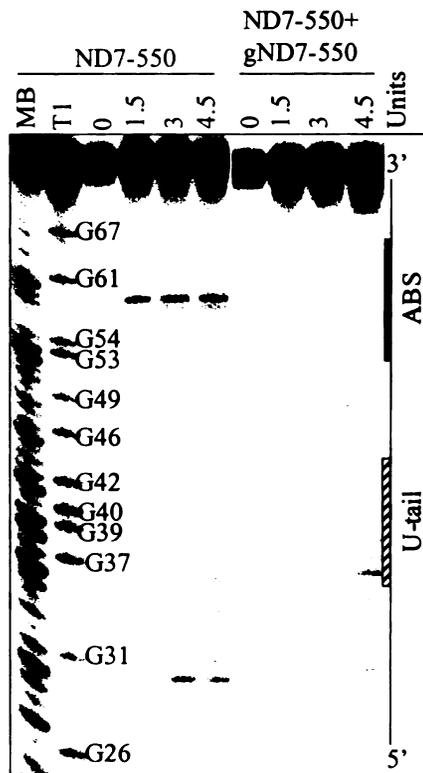


Figure 3.8 Solution Structure Probing of ND7-550/gND7-550. Representative image of a denaturing polyacrylamide gel. The ND7-550 mRNA was 5'-end-labeled and hybridized alone or with gND7-550. 0: no enzyme control. 1.5, 3, and 4.5 Units: amount of Mung Bean nuclease. MB and T1 are Mung Bean and RNase T1 digests for sequence mapping. The position of the ABS and the region where the gRNA U-tail is predicted to bind on the mRNA are indicated.

## DISCUSSION

More than a thousand gRNAs are necessary to generate mature, translatable transcripts of the mitochondrial mRNAs (Souza et al. 1997; Stuart et al. 2005; Ochsenreiter et al. 2007). Although the control mechanism for editing is unknown, editing is developmentally regulated and is not initiated until the gRNA anchor anneals with the mRNA anchor-binding site (Seiwert et al. 1996). Despite the crucial role that gRNAs play in the editing process, very little is known about how the gRNAs specifically and efficiently target their cognate mRNAs. Extensive work has been conducted to characterize the base composition of editing sites (Burgess and Stuart 2000). Mutations, deletions, substitutions, and even a more detailed selection-amplification technique have been applied to regions flanking editing sites to investigate what sequences or structural motifs define such specific regions (Cruz-Reyes et al. 2001; Kabb et al. 2001; Igo et al.

2002; Pai et al. 2003; Golden and Hajduk 2006; Cifuentes-Rojas et al. 2007). However, while these experiments have defined the determinants that “enhance” the efficiency of the in vitro editing reaction, it is unclear if the introduced changes affect protein recognition sites, or simply the ability of the gRNA to target and correctly pair with the selected editing substrate. Considering the existence of hundreds of anchor binding sites within 12 different mitochondrial mRNAs, there must be a variety of secondary and tertiary features that must be disrupted for gRNAs to bind the ABS. The sequence and structure surrounding the ABS could be influencing the nucleation event by the gRNA and also the binding affinity of this bimolecular interaction. Once the duplex between ABS and gRNA anchor forms, it should be particularly stable to allow further reorganization of the two RNAs into the core structure necessary for proper editing (Leung and Koslowsky 2001a; Cifuentes-Rojas et al. 2007).

The results of this work clearly show the impact of mRNA structure on gRNA targeting. Through EMSA and SPR studies we found strong affinities coupled with fast association for the A6U and ND7-550 pairs. The ND7-550/gND7-550 interaction was tighter than the A6U/gA6-14 interaction ( $K_D$  of ~0.5 nM for ND7-550/gND7-550 and ~2.6 nM for A6U/gA6-14). That difference was due to a five fold faster association rate constant, which was surprising as we observed a much faster RNase H ODN directed cleavage of the A6 mRNA in the accessibility assays. This may be readily explained by the fact that the ODN is a short piece of DNA and interaction with its target may not require the intermolecular rearrangements necessary for the larger gRNA interaction. The difference in the A6 and ND7-550 SPR-measured association rate could also be correlated with the G-C content of both anchor-binding sites. The ND7-550 ABS

contains 7 G-C pairs versus only 4 in the A6 (Fig. 3.1). The rate-limiting step in most RNA/RNA interactions is an initial base-pairing interaction that forms in a concentration dependent second-order process. The large number of possible G-C pairs may allow a more stable nucleus to form, increasing the probability of the initial interaction continuing on to helix formation. The most accessible mRNAs, A6 and ND7-550 were also shown to be cleaved by a gRNA-dependent endonuclease activity present in the mitochondrial extract. In contrast to the A6 mRNA, which was specifically cleaved at a single site, 1 nucleotide upstream of the anchor duplex, the ND7-550 was cleaved at 5 sites that coincided with upstream editing sites. Cleavages upstream of the first editing site have been previously reported to happen *in vitro* and *in vivo*, and are of unknown causes (Decker and Sollner-Webb 1990; Kable et al. 1996; Adler and Hajduk 1997). Adler and Hajduk (1997) speculated reasons for multiple cleavages upstream of the expected ES, including improper assembly of the exogenous mRNA/gRNA with the purified editing complexes. More recent studies found that the structure surrounding the ES was a strong determinant of association and cleavage by purified editing complexes (Cifuentes-Rojas et al. 2005; Cifuentes-Rojas et al. 2007; Hernandez et al. 2008). Therefore, structural differences around the ES between the ND7-550 complex and A6 may reflect the distinct cleavage pattern. The ND7-550 substrate differs from A6 in that it has several C-residues located within the first 10 nt 5' of the anchor (Fig. 3.8). Thus, the U-tail interaction is predicted to be pushed further upstream. It may be that this limits the gRNA's ability to direct editing to the 3' most site. Our experiments indicate that the region between the ABS and the sequence where the U-tail is predicted to bind is protected against digestion by a single stranded specific nuclease. This indicates that the

cleavages observed in the ND7-550 were not simply due to “nonspecific” nucleases cleaving single-stranded “hotspots”. Instead, the gRNA appears to be positioning the editing sites within the ND7-550 mRNA in a way that the endoribonuclease present in the editing complex cleaves most often at C4. It may be that while A6 can present the correct editing site to the editosome in the absence of additional proteins (proteins not found in the core editosome), other substrates, like ND7-550, may need additional accessory factors for proper folding and correct presentation of the editing site. Alternatively, it may be that site selection is dictated by thermodynamic stabilities between the gRNA and mRNA, with less favorable sites being edited later (Koslowsky et al. 1991; Alatortsev et al. 2008).

In contrast to both A6 and ND7-550, the two substrates with anchor-binding sites found within highly structured region were not able to efficiently pair with their gRNAs. These targets would probably require accessory proteins for efficient editing. A large number of putative accessory editing proteins have been identified. These include the MRP complex and RBP16, both of which have been shown to have RNA annealing or chaperone-type activities (Muller and Goring 2002; Aphasizhev et al. 2003; Ammerman et al. 2008; Zikova et al. 2008). Other mitochondrial RNA-binding proteins that may also be accessory proteins include mHel61p, REAP-1, TbRGG1, and TbRGG2 (Missel et al. 1997; Hans et al. 2007; Fisk et al. 2008; Hashimi et al. 2008). The MRP complex and RBP16 appear to help modulate editing of specific mRNAs by facilitating mRNA/gRNA annealing. In vitro, annealing of the A6 pair is improved in the presence of MRP (Muller et al. 2001; Zikova et al. 2008) or the RBP16 (Ammerman et al. 2008). However, the A6 mRNA does not require these proteins to be edited in vivo, as

determined by RNAi studies (Pelletier and Read 2003; Vondruskova et al. 2005; Zikova et al. 2006). Disruption of gBP21 expression (from the MRP complex) caused no changes on the levels of edited A6 mRNA, and it actually increased the levels of the edited ND7 message (Vondruskova et al. 2005). Similar conflicting data is also observed with RBP16. RBP16 has been shown to increase the annealing of mutant CYb and A6 pairs in vitro (Miller et al. 2006) but its knockdown only decreased the levels of the edited CYb mRNA and not A6 (Pelletier and Read 2003). The role of these or other proteins in facilitating editing is likely to be transcript-specific, as both have been shown to increase the levels of particular edited mRNAs. This suggests that some mRNA/gRNA pairs, like the A6 substrate, do not need the help of chaperones or matchmakers due to their ability to efficiently hybridize and correctly cleave their target. Some substrates, like ND7-550, may efficiently target the mRNA, yet require an accessory protein for proper folding and presentation of the correct editing sites. Others, like CYb and ND7-506, may require multiple accessory proteins for efficient editing.

Considering the hundreds of different anchor binding sites, it seems plausible that gRNA targeting will be specific and will require different sets of accessory factors, including annealing proteins, proteins that assist proper folding or correct possible misfoldings and/or proteins that bind and stabilize active structures. Identifying substrate pairs with distinct interaction requirements may aid in defining the roles of the different accessory proteins.

## MATERIALS AND METHODS

### Oligonucleotides

All ODNs (Table 3.2) were purchased from Integrated DNA Technologies, Inc.

(Coralville, IA). The oligoribonucleotide ND7-550 was obtained from Dharmacon

(Boulder, CO) for the SPR experiments: 5'-AAAAACAUGACUACAUGAUUAGUACAA

GAGGAGACAGACGACAGUGUCCACAGCACCCGUUUCAGCACAG-3'.

Table 3.2 Oligodeoxyribonucleotides.

ODN name	Sequence (5' to 3')
<b>gRNA</b>	
T7-22	AATTTAATACGACTCACTATAG
gA6-14	AAAAAAAAAAAAAAAAAATAATTATCATATCACTGTCAAAAATCTGATT CGTTATCGGAGTTATAGCCCTATAGTGAGTCGTATTAAT
gND7-550	AAAAAAAAAAAAAAAAAATAATTCACATTTATATCATCTTACACTTAATC CACTGCATCCCTATAGTGAGTCGTATTAAT
gND7-550sU	TATTCACATTTATATCATCTTACACTTAATCCACTGCATCCCTATAG TGAGTCGTATTAAT
gND7-506	AAAAAAAAAAAAAAAAAATTCACTATCTACACTAACTATACTACAGGT TATTTACATCGTAGCCCTATAGTGAGTCGTATTAAT
gCYb-558	AAAAAAAAAAAAAAAAAATTATCCCTTTATCACCTAGAAATTCACAT TGCTTTTAATCCCTATAGTGAGTCGTATTAAT
<b>mRNA</b>	
ND7UHR3 - 5' half	CATCAATAAATCTTATCCCCTCTCCTCCAACGTGCCTATAGTGAG TCGTATTAAT
ND7UHR3 - 3' half	CATTGTTCTACACTTTTATATTACATAACTTTTCTGTACCACGATG CAAATCACAAATTT
ND7UHR3 Bridge	GATAAGATTTATTGATGAAATTTGTGATTTGC
ND7UHR3 Forward	AATTTAATACGACTCACTATAG
ND7UHR3 Reverse	CATTGTTCTACACTTTTATATTACATAAC
ND7-550 - 5' half	AATTTAATACGACTCACTATAGGGATACAAAAAACATG
ND7-550 - 3' half	ACTACATGATAAGTACAAGAGGAGACAGACGACAGGTCCACAGC ACCCGTTTCA
ND7-550 Bridge	TACTTATCATGTAGTCATGTTTTTTGTATC
ND7-550 Forward	AATTTAATACGACTCACTATAGGGATACAAAAAACATG
ND7-550 Reverse	GTGCTGAAACGGGTGCTGTGGACCTGTCGTC
T7A6 Forward	AATTTAATACGACTCACTATAGGAAAGG
A6U Reverse	TATTATTAACCTATTTGATCTTATTCTATAACTCCAA
A6P1 Reverse	mUmATTTGATCTTATTCTATAACTCCAATCACAAAC
A6U Reverse	CTTATTTGATCTTATTCTATAACTCCAA
T7CYb Forward	AATTTAATACGACTCACTATAGGGTTATAAAT
CYb Reverse	GGCCGCTCTAGAACTAGTGG
<b>RNase H assay ODNs</b>	
A6U	CTATAACTCC
CYbU	ATTAAGAGAC
ND7UHR3	CGATGCAAATC
ND7-550	GTGCTGTGGAC
<b>SPR assays</b>	
3'BigSK-biotin	CACTAGTTCTAGAGCGGCC-biotin
A6BigskBIAbriidge	GCTCTAGAACTAGTGTATTATTAACCTATTTG
ND7550BigskBIAbriidge	GCTCTAGAACTAGTGTGTGCTGAAACGGG

### **Templates for RNA transcription**

A6 and CYb templates were PCR amplified using the forward and reverse primers listed above and plasmids described previously (Koslowsky et al. 1996). Templates for ND7UHR3 and ND7-550 were generated by ligation of 1 nmol of 5' <sup>32</sup>P-labeled 3' half ODN to 1 nmol of the 5' half ODN listed above, using 1 nmol of the bridge ODN and 25 U of T4 DNA Ligase (Roche) in 66 mM Tris-HCl, 5 mM MgCl<sub>2</sub>, 5 mM DTT, 1 mM ATP, pH 7.5, at 22°C, overnight (Moore and Sharp 1992). The ligated single stranded DNA product was gel purified on 8% (w/v) 8M urea polyacrylamide gel and then amplified using the appropriate forward and reverse primers and Taq DNA polymerase (Promega) as recommended by the manufacturer.

### **RNA transcription and radioactive 5'-end-labeling**

RNAs were transcribed using the T7 RiboMax kit (Promega) according to manufacturer directions. For the end-labeling, the 5' phosphates from the RNAs were removed using Calf Intestinal Alkaline Phosphatase (New England Biolabs) and further 5'-end-labeled with 50 μCi of [ $\gamma$ <sup>32</sup>P]-ATP, using T4 Polynucleotide Kinase (Invitrogen) and standard procedures. The transcribed or labeled RNAs were gel purified by electrophoresis on 8% (mRNA) or 15% (gRNA) polyacrylamide gels containing 8M urea. Bands corresponding to desired RNAs were localized on the gels by UV shadowing or autoradiography. The RNAs were eluted in an RNA elution buffer (10 mM Tris pH 7.8, 0.1% SDS, 2 mM EDTA, and 0.3 M NaOAc pH 7.0) in the presence of phenol, recovered by ethanol precipitation, suspended in RNase-free H<sub>2</sub>O, and quantified using a Cary 50 spectrophotometer.

## **Secondary Structure Prediction**

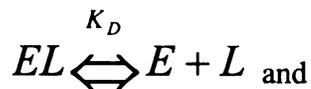
Predicted secondary structures and free energies were obtained using mfold version 2.3, <http://mfold.bioinfo.rpi.edu/> (Walter et al. 1994; Zuker 2003) and DINAMelt, <http://dinamelt.bioinfo.rpi.edu/> (Markham and Zuker 2005). Confirmation of the predicted structures and accessibility of the ABS was obtained using ODN directed RNase H assays.

## **ODN directed RNase H assays**

Fifty picomoles of 5' <sup>32</sup>P-labeled mRNAs were renatured after gel-purification by heating to 70°C for 3 min and slow cooling (2°C/min) to 27°C in RH buffer (40 mM Tris-HCl pH 7.5, 100 mM KCl, 2 mM MgCl<sub>2</sub>, 1 mM DTT). The sample was incubated at 27°C for 30 additional min and then quenched on ice. One unit of RNase H and different concentrations of the appropriate ODN (at the molar ratios of 1:1, 1:5, 1:10, and 1:30 substrate:ODN) were added. From the 50 µl reaction, an initial 10 µl aliquot was immediately taken (after about 1 min of incubation). The reaction was returned to 27°C, and further aliquots taken at 15 and 30 min. Reactions were stopped with addition of formamide loading buffer (80% (v/v) formamide, 10 mM EDTA, 1 mg/ml bromophenol blue, 1 mg/ml xylene cyanol) and quenched on ice. Samples were solved on 8% 8M Urea polyacrylamide gels. The gels were fixed, dried, and exposed on a PhosphorImaging screen overnight. All assays were conducted in triplicate. The percentage of RNase H digestion (radioactive bands) was determined using a Storm PhosphorImager (Molecular Dynamics, Sunnyvale, CA). The fraction cleaved was calculated as the signal in the band corresponding to cleaved mRNA divided by the total signal of the cleaved and free bands.

## Determination of Binding Affinity

Binding affinities were determined by Electrophoretic Mobility Shift Assays (EMSA) as previously described (Koslowsky et al. 2004) except that the gRNA/mRNA hybridization time was increased to 3 hours. The apparent affinity constant ( $K_D$ ) of gRNA binding was extracted from data-point fitting using KaleidaGraph 3.5 and the following binding model (Matthews 1993):



**Equation 3.1 Dissociation equilibrium constant**

$$K_D = \frac{[L_f][E_f]}{[EL]}$$

$$[L_f] = \text{free ligand} = [L_T] - [EL]$$

$$[E_f] = \text{free binding sites} = [E_T] - [EL]$$

$$[E_T] = \text{Total concentration of binding sites} = n[M]$$

$$[M] = \text{mRNA concentration}$$

$$n = \text{number of binding sites}$$

In our case,  $n = 1$ , thus  $[E_f] = [M_f]$ , or:

**Equation 3.2 Dissociation equilibrium constant for the mRNA/gRNA complex**

$$K_D = \frac{[gRNA_f][mRNA_f]}{[complex]}$$

Where: complex = gRNA bound to mRNA

$$[gRNA_{free}] = [gRNA_{total}] - [complex] \text{ and}$$

$$[mRNA_{free}] = [mRNA_{total}] - [complex]$$

The  $K_D$  value is an average of four experiments and the standard deviation was calculated from the difference in these values. The A6U mRNA used for the gel shifts was shorter at the 3' end by 9 nucleotides, a sequence that has been shown not to interfere with editing efficiency (Cifuentes-Rojas et al. 2005; Hernandez et al. 2008).

### **Determination of rate constants**

The mRNA/gRNA pairs used in this study were A6U/gA6-14 and ND7-550/gND7-550, described above. These substrates were all synthesized in vitro with the exception of ND7-550 that was chemically synthesized by Dharmacon Inc. The running or binding buffer utilized for these studies was 100 mM Tris pH 7.5, 0.1 mM EDTA, 2 mM  $MgCl_2$ , and 100 mM KCl and the regeneration buffer was 8 M Urea. Measurements of the association and dissociation rate constants were performed on a BIACORE 2000 (BIACORE, Uppsala, Sweden). All the solutions used in the binding studies were filtered through a 0.22  $\mu m$  polyethersulfone membrane (Corning) or a 0.22  $\mu m$  Millex-GS membrane (Millipore) and degassed. The mRNAs were ligated to the 3' BigSK-biotin ODN tag (IDT DNA Technologies, Inc.) by annealing the tag to the 3' end of the appropriate mRNA using a bridging ODN as previously described (Yu and Koslowsky 2006). The biotinylated mRNAs were gel extracted without phenol and purified using ultra-free MC membranes (Millipore) and microcon tubes (YM-50, Millipore) according to the manufacturer's directions. The gRNAs were also purified using Ultra-free MC and the YM-10 or 30 microcon tubes. The RNA samples were then diluted in running buffer. The biotinylated mRNA was diluted to 10 nM and 50-400 resonance units (RU) of mRNA was attached at 5  $\mu l/min$  to a streptavidin coated SA sensor chip (BIACORE, Uppsala, Sweden); the better the mRNA/gRNA interaction the less mRNA was attached.

Two cells were immobilized with mRNA, one was left unmodified to serve as a reference cell, and one cell was immobilized with ODN tag as a control cell. Binding studies were carried out running all four cells in series with respective cycles: 1) 100-300  $\mu\text{l}$  gRNA injection at 5-10  $\mu\text{l}/\text{min}$  (to obtain association at varying concentrations of gRNA). 2) Buffer flow (for dissociation of gRNA) 5-10  $\mu\text{l}/\text{min}$  for 15-60 min. 3) Regeneration (50  $\mu\text{l}$  injection of regeneration buffer, two 50  $\mu\text{l}$  injections of running buffer at 50  $\mu\text{l}/\text{min}$ ). These experiments were conducted at 27°C. The determination of rate constants was performed by fitting theoretical curves to the experimental curves using BIAevaluation 3.0 software (BIACORE). The equation used to calculate the association rate constant from the Biacore sensorgrams was the 1:1 (Langmuir) association formula describing analyte (gRNA) binding to ligand (mRNA):

**Equation 3.3 Association rate constant**

$$\frac{k_a * Conc * R_{\max}}{(k_a * Conc + k_d)} \times \left(1 - e^{-(k_a * Conc + k_d) * (t - t_0)}\right) + RI$$

$k_a$  = association rate constant ( $k_{\text{on}}$ ),  $Conc$  = molar analyte concentration,  $RI$  = bulk refractive index effect (RU),  $R_{\max} = Y_{\max}$  = maximum analyte binding capacity (RU),  $t$  = time,  $t_0$  = time at start,  $k_d$  = dissociation rate constant ( $k_{\text{off}}$ ). The equation used to calculate the dissociation rate constant was the 1:1 (Langmuir) dissociation formula describing analyte (gRNA) dissociating from surface complex (mRNA/gRNA):

**Equation 3.4 Dissociation rate constant**

$$R_0 * e^{(-k_d * (t - t_0))} + Offset$$

$RO = Y_{\max}$  = maximum analyte binding capacity (RU),  $k_d$  = dissociation rate constant ( $k_{\text{off}}$ ),  $t$  = time,  $t_0$  = time at start, *Offset* = residual response at infinite time (RU).

Separate fits for each association and dissociation curves were analyzed globally from each experiment to obtain  $k_{\text{on}}$  and  $k_{\text{off}}$ , individually, and the results were averaged. The dissociation equilibrium constant ( $K_D$ ) was calculated from the averages of the rate constants using the equation:

**Equation 3.5 Dissociation equilibrium constant**

$$K_D = \frac{k_d}{k_a}$$

The standard deviation reported for the rate constants were based on the variances of all curves (Nordgren et al. 2001).

### **gRNA Directed Cleavage Assays**

Procyclic-form *T. brucei* were grown and the mitochondria were isolated, lysed and cleared by centrifugation. The trypanosome mitochondrial extract ( $\sim 2 \times 10^{10}$  cell equivalents/mL) was then separated by glycerol gradient sedimentation as previously described (Pollard et al. 1992; Seiwert et al. 1996; Leung and Koslowsky 2001a). We additionally used a mitochondrial extract enriched for the editing complex by Q-Sepharose chromatography, obtained from Dr. Cruz-Reyes (Rusche et al. 1997; Cruz-Reyes et al. 2002). Both purified editing complexes were further treated with 0.5 mM inorganic pyrophosphate (PPi) to inhibit the RNA ligase reaction and improve the cleavage efficiency (Cruz-Reyes et al. 1998). The gRNA directed cleavage assays were conducted in triplicate with editing complexes from the glycerol fractions and Cruz-Reyes' purified complexes. For each cleavage reaction, approximately 0.1 pmols of 5'-

<sup>32</sup>P-labeled mRNA (60 Kcpm) and 1 pmol of cognate gRNA were heated to 70°C (3 min), slow cooled (2°C/min) to 27°C and incubated at 27°C for 30 min. Then, 2 µl of PPI-treated purified editing complexes or 10 µl of the most active glycerol fraction were added and the reaction incubated for an additional hour in 10 mM KCl-MRB buffer (25 mM Tris-HCl pH 7.9, 10 mM MgOAc, 10 mM KCl, 1 mM EDTA, 0.5 mM DTT, 1 mM CaCl<sub>2</sub>, 5% glycerol). The cleavage reaction was terminated by the addition of 2 µl of stop buffer (130 mM EDTA, 2.5% SDS) followed by phenol/chloroform/isoamyl alcohol (25:24:1) extraction and ethanol precipitation. Samples were dried, suspended in formamide loading buffer, and resolved on 8-12% (w/v) denaturing polyacrylamide gels as described before. The product amount was calculated as the percentage of total input mRNA.

### **Solution Structure Probing**

To probe for single stranded regions in the ND7-550 mRNA when hybridized with gND7-550, we used 25 Kcpm of 5'-<sup>32</sup>P-labeled mRNA for the control reactions without enzyme and 50 Kcpm (~20 fmols) for the other reactions. The gRNA was added exceeding the mRNA concentration by 10 times. Both RNAs were hybridized in NEB buffer 2 (10 mM Tris pH 7.5, 50 mM NaCl, 10 mM MgCl<sub>2</sub>, and 1 mM DTT), by heating to 70°C for 3 min, slowly cooling to 27°C, and keeping at 27°C for approximately 1h. The single stranded specific Mung Bean Nuclease (NEB, Ipswich, MA) was then added (1.5 U, 3 U, or 4.5 U) and the sample was incubated at 27°C for 10 min. The reaction was phenol/chloroform/isoamyl alcohol (25:24:1) extracted and treated as described in the gRNA directed cleavage assays.

## **ACKNOWLEDGEMENTS**

This work was funded by the NIH (AI45835 to D.J.K.). We thank Drs. Ronald Patterson, Charles Hoogstraten, John Wang, Patrick Walton, Robert Britton, and other members of the MSU RNA Journal Club for their helpful comments on the work. We would like to thank Dr. Joseph Leykam and the members of the Macromolecular Structure Facility in the MSU Biochemistry and Molecular Biology Department for the use of the Biacore machine.

## REFERENCES

- Adler, B.K., and Hajduk, S.L. 1997. Guide RNA requirement for editing-site-specific endonucleolytic cleavage of preedited mRNA by mitochondrial ribonucleoprotein particles in *Trypanosoma brucei*. *Mol Cell Biol* **17**: 5377-5385.
- Alatortsev, V.S., Cruz-Reyes, J., Zhelonkina, A.G., and Sollner-Webb, B. 2008. *Trypanosoma brucei* RNA editing: coupled cycles of U deletion reveal processive activity of the editing complex. *Mol Cell Biol* **28**: 2437-2445.
- Amarzguioui, M., Brede, G., Babaie, E., Grotli, M., Sproat, B., and Prydz, H. 2000. Secondary structure prediction and in vitro accessibility of mRNA as tools in the selection of target sites for ribozymes. *Nucleic Acids Res* **28**: 4113-4124.
- Ammerman, M.L., Fisk, J.C., and Read, L.K. 2008. gRNA/pre-mRNA annealing and RNA chaperone activities of RBP16. *RNA* **14**: 1069-1080.
- Aphasizhev, R., Aphasizheva, I., Nelson, R.E., and Simpson, L. 2003. A 100-kD complex of two RNA-binding proteins from mitochondria of *Leishmania tarentolae* catalyzes RNA annealing and interacts with several RNA editing components. *RNA* **9**: 62-76.
- Argaman, L., Hershberg, R., Vogel, J., Bejerano, G., Wagner, E.G., Margalit, H., and Altuvia, S. 2001. Novel small RNA-encoding genes in the intergenic regions of *Escherichia coli*. *Curr Biol* **11**: 941-950.
- Benne, R. 1994. RNA editing in trypanosomes. *Eur J Biochem* **221**: 9-23.
- Benne, R., Van den Burg, J., Brakenhoff, J.P., Sloof, P., Van Boom, J.H., and Tromp, M.C. 1986. Major transcript of the frameshifted coxII gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA. *Cell* **46**: 819-826.
- Birikh, K.R., Berlin, Y.A., Soreq, H., and Eckstein, F. 1997. Probing accessible sites for ribozymes on human acetylcholinesterase RNA. *RNA* **3**: 429-437.
- Blum, B., Bakalara, N., and Simpson, L. 1990. A model for RNA editing in kinetoplastid mitochondria: "guide" RNA molecules transcribed from maxicircle DNA provide the edited information. *Cell* **60**: 189-198.
- Blum, B., and Simpson, L. 1990. Guide RNAs in kinetoplastid mitochondria have a nonencoded 3' oligo(U) tail involved in recognition of the preedited region. *Cell* **62**: 391-397.
- Blum, B., and Simpson, L. 1992. Formation of guide RNA/messenger RNA chimeric molecules in vitro, the initial step of RNA editing, is dependent on an anchor sequence. *Proc Natl Acad Sci U S A* **89**: 11944-11948.

- Brennecke, J., Stark, A., Russell, R.B., and Cohen, S.M. 2005. Principles of microRNA-target recognition. *PLoS Biol* **3**: e85.
- Brown, K.M., Chu, C.Y., and Rana, T.M. 2005. Target accessibility dictates the potency of human RISC. *Nat Struct Mol Biol* **12**: 469-470.
- Burgess, M.L., and Stuart, K. 2000. Sequence bias in edited kinetoplastid RNAs. *RNA* **6**: 1492-1497.
- Campbell, T.B., McDonald, C.K., and Hagen, M. 1997. The effect of structure in a long target RNA on ribozyme cleavage efficiency. *Nucleic Acids Res* **25**: 4985-4993.
- Carnes, J., Trotter, J.R., Peltan, A., Fleck, M., and Stuart, K. 2008. RNA editing in *Trypanosoma brucei* requires three different editosomes. *Mol Cell Biol* **28**: 122-130.
- Cifuentes-Rojas, C., Halbig, K., Sacharidou, A., De Nova-Ocampo, M., and Cruz-Reyes, J. 2005. Minimal pre-mRNA substrates with natural and converted sites for full-round U insertion and U deletion RNA editing in trypanosomes. *Nucleic Acids Res* **33**: 6610-6620.
- Cifuentes-Rojas, C., Pavia, P., Hernandez, A., Osterwisch, D., Puerta, C., and Cruz-Reyes, J. 2007. Substrate determinants for RNA editing and editing complex interactions at a site for full-round U insertion. *J Biol Chem* **282**: 4265-4276.
- Cruz-Reyes, J., Rusche, L.N., Piller, K.J., and Sollner-Webb, B. 1998. *T. brucei* RNA editing: adenosine nucleotides inversely affect U-deletion and U-insertion reactions at mRNA cleavage. *Mol Cell* **1**: 401-409.
- Cruz-Reyes, J., Zhelonkina, A., Rusche, L., and Sollner-Webb, B. 2001. Trypanosome RNA editing: simple guide RNA features enhance U deletion 100-fold. *Mol Cell Biol* **21**: 884-892.
- Cruz-Reyes, J., Zhelonkina, A.G., Huang, C.E., and Sollner-Webb, B. 2002. Distinct functions of two RNA ligases in active *Trypanosoma brucei* RNA editing complexes. *Mol Cell Biol* **22**: 4652-4660.
- Decker, C.J., and Sollner-Webb, B. 1990. RNA editing involves indiscriminate U changes throughout precisely defined editing domains. *Cell* **61**: 1001-1011.
- Eckardt, S., Romby, P., and Sczakiel, G. 1997. Implications of RNA structure on the annealing of a potent antisense RNA directed against the human immunodeficiency virus type 1. *Biochemistry* **36**: 12711-12721.
- Feagin, J.E., Jasmer, D.P., and Stuart, K. 1987. Developmentally regulated addition of nucleotides within apocytochrome b transcripts in *Trypanosoma brucei*. *Cell* **49**: 337-345.

- Fisk, J.C., Ammerman, M.L., Presnyak, V., and Read, L.K. 2008. TbRGG2, an essential RNA editing accessory factor in two *Trypanosoma brucei* life cycle stages. *J Biol Chem* **283**: 23016-23025.
- Franch, T., Petersen, M., Wagner, E.G., Jacobsen, J.P., and Gerdes, K. 1999. Antisense RNA regulation in prokaryotes: rapid RNA/RNA interaction facilitated by a general U-turn loop structure. *J Mol Biol* **294**: 1115-1125.
- Golden, D.E., and Hajduk, S.L. 2006. The importance of RNA structure in RNA editing and a potential proofreading mechanism for correct guide RNA:pre-mRNA binary complex formation. *J Mol Biol* **359**: 585-596.
- Hackermuller, J., Meisner, N.C., Auer, M., Jaritz, M., and Stadler, P.F. 2005. The effect of RNA secondary structures on RNA-ligand binding and the modifier RNA mechanism: a quantitative model. *Gene* **345**: 3-12.
- Hans, J., Hajduk, S.L., and Madison-Antenucci, S. 2007. RNA-editing-associated protein 1 null mutant reveals link to mitochondrial RNA stability. *RNA* **13**: 881-889.
- Hashimi, H., Zikova, A., Panigrahi, A.K., Stuart, K.D., and Lukes, J. 2008. TbRGG1, an essential protein involved in kinetoplastid RNA metabolism that is associated with a novel multiprotein complex. *RNA* **14**: 970-980.
- Hermann, T., Schmid, B., Heumann, H., and Goring, H.U. 1997. A three-dimensional working model for a guide RNA from *Trypanosoma brucei*. *Nucleic Acids Res* **25**: 2311-2318.
- Hernandez, A., Panigrahi, A., Cifuentes-Rojas, C., Sacharidou, A., Stuart, K., and Cruz-Reyes, J. 2008. Determinants for association and guide RNA-directed endonuclease cleavage by purified RNA editing complexes from *Trypanosoma brucei*. *J Mol Biol* **381**: 35-48.
- Igo, R.P., Jr., Lawson, S.D., and Stuart, K. 2002. RNA sequence and base pairing effects on insertion editing in *Trypanosoma brucei*. *Mol Cell Biol* **22**: 1567-1576.
- Kabb, A.L., Opegard, L.M., McKenzie, B.A., and Connell, G.J. 2001. A mRNA determinant of gRNA-directed kinetoplastid editing. *Nucleic Acids Res* **29**: 2575-2580.
- Kable, M.L., Seiwert, S.D., Heidmann, S., and Stuart, K. 1996. RNA editing: a mechanism for gRNA-specified uridylyate insertion into precursor mRNA. *Science* **273**: 1189-1195.
- Katsamba, P.S., Park, S., and Laird-Offringa, I.A. 2002. Kinetic studies of RNA-protein interactions using surface plasmon resonance. *Methods* **26**: 95-104.
- Kertesz, M., Iovino, N., Unnerstall, U., Gaul, U., and Segal, E. 2007. The role of site accessibility in microRNA target recognition. *Nat Genet* **39**: 1278-1284.

- Kikuchi, K., Umehara, T., Fukuda, K., Hwang, J., Kuno, A., Hasegawa, T., and Nishikawa, S. 2003. RNA aptamers targeted to domain II of hepatitis C virus IRES that bind to its apical loop region. *J Biochem* **133**: 263-270.
- Koslowsky, D.J., Bhat, G.J., Perrollaz, A.L., Feagin, J.E., and Stuart, K. 1990. The MURF3 gene of *T. brucei* contains multiple domains of extensive editing and is homologous to a subunit of NADH dehydrogenase. *Cell* **62**: 901-911.
- Koslowsky, D.J., Bhat, G.J., Read, L.K., and Stuart, K. 1991. Cycles of progressive realignment of gRNA with mRNA in RNA editing. *Cell* **67**: 537-546.
- Koslowsky, D.J., Kutas, S.M., and Stuart, K. 1996. Distinct differences in the requirements for ribonucleoprotein complex formation on differentially regulated pre-edited mRNAs in *Trypanosoma brucei*. *Mol Biochem Parasitol* **80**: 1-14.
- Koslowsky, D.J., Reifur, L., Yu, L.E., and Chen, W. 2004. Evidence for U-tail stabilization of gRNA/mRNA interactions in kinetoplastid RNA editing. *RNA Biol* **1**: 28-34.
- Lawson, S.D., Igo, R.P., Jr., Salavati, R., and Stuart, K.D. 2001. The specificity of nucleotide removal during RNA editing in *Trypanosoma brucei*. *RNA* **7**: 1793-1802.
- Leung, S.S., and Koslowsky, D.J. 2001a. Interactions of mRNAs and gRNAs involved in trypanosome mitochondrial RNA editing: structure probing of an mRNA bound to its cognate gRNA. *RNA* **7**: 1803-1816.
- Lewis, B.P., Burge, C.B., and Bartel, D.P. 2005. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **120**: 15-20.
- Lima, W.F., Monia, B.P., Ecker, D.J., and Freier, S.M. 1992. Implication of RNA structure on antisense oligonucleotide hybridization kinetics. *Biochemistry* **31**: 12055-12061.
- Long, D., Chan, C.Y., and Ding, Y. 2008. Analysis of microRNA-target interactions by a target structure based hybridization model. *Pac Symp Biocomput*: 64-74.
- Long, D., Lee, R., Williams, P., Chan, C.Y., Ambros, V., and Ding, Y. 2007. Potent effect of target structure on microRNA function. *Nat Struct Mol Biol* **14**: 287-294.
- Madison-Antenucci, S., Grams, J., and Hajduk, S.L. 2002. Editing machines: the complexities of trypanosome RNA editing. *Cell* **108**: 435-438.
- Markham, N.R., and Zuker, M. 2005. DINAMelt web server for nucleic acid melting prediction. *Nucleic Acids Res* **33**: W577-581.
- Matthews, J. 1993. *Simple receptor-ligand interactions. Fundamentals of Receptor, Enzyme and Transport Kinetics*. CRC Press Inc., pp. 7-23.

- Mercatanti, A., Rainaldi, G., Mariani, L., Marangoni, R., and Citti, L. 2002. A method for prediction of accessible sites on an mRNA sequence for target selection of hammerhead ribozymes. *J Comput Biol* **9**: 641-653.
- Miller, M.M., Halbig, K., Cruz-Reyes, J., and Read, L.K. 2006. RBP16 stimulates trypanosome RNA editing in vitro at an early step in the editing reaction. *RNA* **12**: 1292-1303.
- Missel, A., Souza, A.E., Norskau, G., and Goringe, H.U. 1997. Disruption of a gene encoding a novel mitochondrial DEAD-box protein in *Trypanosoma brucei* affects edited mRNAs. *Mol Cell Biol* **17**: 4895-4903.
- Moore, M.J., and Sharp, P.A. 1992. Site-specific modification of pre-mRNA: the 2'-hydroxyl groups at the splice sites. *Science* **256**: 992-997.
- Muckstein, U., Tafer, H., Hackermuller, J., Bernhart, S.H., Stadler, P.F., and Hofacker, I.L. 2006. Thermodynamics of RNA-RNA binding. *Bioinformatics* **22**: 1177-1182.
- Muller, U.F., and Goringe, H.U. 2002. Mechanism of the gBP21-mediated RNA/RNA annealing reaction: matchmaking and charge reduction. *Nucleic Acids Res* **30**: 447-455.
- Muller, U.F., Lambert, L., and Goringe, H.U. 2001. Annealing of RNA editing substrates facilitated by guide RNA-binding protein gBP21. *Embo J* **20**: 1394-1404.
- Myszka, D.G. 1997. Kinetic analysis of macromolecular interactions using surface plasmon resonance biosensors. *Curr Opin Biotechnol* **8**: 50-57.
- Nordgren, S., Slagter-Jager, J.G., and Wagner, G.H. 2001. Real time kinetic studies of the interaction between folded antisense and target RNAs using surface plasmon resonance. *J Mol Biol* **310**: 1125-1134.
- Ochsenreiter, T., Cipriano, M., and Hajduk, S.L. 2007. KISS: the kinetoplastid RNA editing sequence search tool. *RNA* **13**: 1-4.
- Ochsenreiter, T., Cipriano, M., and Hajduk, S.L. 2008. Alternative mRNA editing in trypanosomes is extensive and may contribute to mitochondrial protein diversity. *PLoS ONE* **3**: e1566.
- Overhoff, M., Alken, M., Far, R.K., Lemaitre, M., Lebleu, B., Sczakiel, G., and Robbins, I. 2005. Local RNA target structure influences siRNA efficacy: a systematic global analysis. *J Mol Biol* **348**: 871-881.
- Pai, R.D., Opegard, L.M., and Connell, G.J. 2003. Sequence and structural requirements for optimal guide RNA-directed insertional editing within *Leishmania tarentolae*. *RNA* **9**: 469-483.

- Patzel, V., and Sczakiel, G. 1998. Theoretical design of antisense RNA structures substantially improves annealing kinetics and efficacy in human cells. *Nat Biotechnol* **16**: 64-68.
- Patzel, V., Steidl, U., Kronenwett, R., Haas, R., and Sczakiel, G. 1999. A theoretical approach to select effective antisense oligodeoxyribonucleotides at high statistical probability. *Nucleic Acids Res* **27**: 4328-4334.
- Pelletier, M., and Read, L.K. 2003. RBP16 is a multifunctional gene regulatory protein involved in editing and stabilization of specific mitochondrial mRNAs in *Trypanosoma brucei*. *RNA* **9**: 457-468.
- Pollard, V.W., Harris, M.E., and Hajduk, S.L. 1992. Native mRNA editing complexes from *Trypanosoma brucei* mitochondria. *Embo J* **11**: 4429-4438.
- Reifur, L., and Koslowsky, D.J. 2008. *Trypanosoma brucei* ATPase subunit 6 mRNA bound to gA6-14 forms a conserved three-helical structure. *RNA* **14**: 2195-2211.
- Rusche, L.N., Cruz-Reyes, J., Piller, K.J., and Sollner-Webb, B. 1997. Purification of a functional enzymatic editing complex from *Trypanosoma brucei* mitochondria. *Embo J* **16**: 4069-4081.
- Schmid, B., Riley, G.R., Stuart, K., and Goring, H.U. 1995. The secondary structure of guide RNA molecules from *Trypanosoma brucei*. *Nucleic Acids Res* **23**: 3093-3102.
- Seiwert, S.D., Heidmann, S., and Stuart, K. 1996. Direct visualization of uridylyate deletion in vitro suggests a mechanism for kinetoplastid RNA editing. *Cell* **84**: 831-841.
- Seiwert, S.D., and Stuart, K. 1994. RNA editing: transfer of genetic information from gRNA to precursor mRNA in vitro. *Science* **266**: 114-117.
- Shao, Y., Chan, C.Y., Maliyekkel, A., Lawrence, C.E., Roninson, I.B., and Ding, Y. 2007a. Effect of target secondary structure on RNAi efficiency. *RNA* **13**: 1631-1640.
- Simpson, L., Sbicego, S., and Aphasizhev, R. 2003. Uridine insertion/deletion RNA editing in trypanosome mitochondria: a complex business. *RNA* **9**: 265-276.
- Souza, A.E., Hermann, T., and Goring, H.U. 1997. The guide RNA database. *Nucleic Acids Res* **25**: 104-106.
- Stuart, K.D., Schnauffer, A., Ernst, N.L., and Panigrahi, A.K. 2005. Complex management: RNA editing in trypanosomes. *Trends Biochem Sci* **30**: 97-105.
- Tafer, H., Ameres, S.L., Obermosterer, G., Gebeshuber, C.A., Schroeder, R., Martinez, J., and Hofacker, I.L. 2008. The impact of target site accessibility on the design of effective siRNAs. *Nat Biotechnol* **26**: 578-583.

- Vickers, T.A., Wyatt, J.R., and Freier, S.M. 2000. Effects of RNA secondary structure on cellular antisense activity. *Nucleic Acids Res* **28**: 1340-1347.
- von der Haar, T., and McCarthy, J.E. 2003. Studying the assembly of multicomponent protein and ribonucleoprotein complexes using surface plasmon resonance. *Methods* **29**: 167-174.
- Vondruskova, E., van den Burg, J., Zikova, A., Ernst, N.L., Stuart, K., Benne, R., and Lukes, J. 2005. RNA interference analyses suggest a transcript-specific regulatory role for mitochondrial RNA-binding proteins MRP1 and MRP2 in RNA editing and other RNA processing in *Trypanosoma brucei*. *J Biol Chem* **280**: 2429-2438.
- Wagner, E.G., and Simons, R.W. 1994. Antisense RNA control in bacteria, phages, and plasmids. *Annu Rev Microbiol* **48**: 713-742.
- Walter, A.E., Turner, D.H., Kim, J., Lyttle, M.H., Muller, P., Mathews, D.H., and Zuker, M. 1994. Coaxial stacking of helices enhances binding of oligoribonucleotides and improves predictions of RNA folding. *Proc Natl Acad Sci USA* **91**: 9218-9222.
- Walton, S.P., Stephanopoulos, G.N., Yarmush, M.L., and Roth, C.M. 2002. Thermodynamic and kinetic characterization of antisense oligodeoxynucleotide binding to a structured mRNA. *Biophys J* **82**: 366-377.
- Wassarman, K.M., Zhang, A., and Storz, G. 1999. Small RNAs in *Escherichia coli*. *Trends Microbiol* **7**: 37-45.
- Ying, S.Y., Chang, D.C., and Lin, S.L. 2008. The microRNA (miRNA): overview of the RNA genes that modulate gene function. *Mol Biotechnol* **38**: 257-268.
- Yu, L.E., and Koslowsky, D.J. 2006. Interactions of mRNAs and gRNAs involved in trypanosome mitochondrial RNA editing: structure probing of a gRNA bound to its cognate mRNA. *RNA* **12**: 1050-1060.
- Zikova, A., Horakova, E., Jirku, M., Dunajcikova, P., and Lukes, J. 2006. The effect of down-regulation of mitochondrial RNA-binding proteins MRP1 and MRP2 on respiratory complexes in procyclic *Trypanosoma brucei*. *Mol Biochem Parasitol* **149**: 65-73.
- Zikova, A., Kopečna, J., Schumacher, M.A., Stuart, K., Trantirek, L., and Lukes, J. 2008. Structure and function of the native and recombinant mitochondrial MRP1/MRP2 complex from *Trypanosoma brucei*. *Int J Parasitol* **38**: 901-912.
- Zuker, M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* **31**: 3406-3415.

## Chapter 4

### **U-TAIL INFLUENCE ON THE BINDING THERMODYNAMICS OF ATPASE SUBUNIT 6 MRNA/GRNA IN *TRYPANOSOMA BRUCEI***

## **ABSTRACT**

The annealing of gRNAs with cognate pre-edited or partially edited mRNAs is crucial to RNA editing in kinetoplastid parasites. Although distinct mRNA/gRNA pairs can have distinct binding affinities they hybridize into a similar secondary structure. The structure is three-helical, including an anchor duplex, a gRNA stem-loop and a U-tail helix. The gRNA U-tail exists at the 3' end of all gRNAs and may play multiple roles in complex formation and editing. Deleting the U-tail in gA6-14 slightly decreases the binding affinity to its cognate A6 mRNA and the amount of final edited product. If the U-tail is modified to allow enhanced binding, editing efficiency increases. In this work, we describe the binding thermodynamics of A6/gA6-14 by using isothermal titration calorimetry. We found that complex formation is energetically favorable and enthalpically driven. Deleting the U-tail results in less negative changes in enthalpy and entropy whereas the enhanced U-tail promotes more negative changes in enthalpy and entropy upon binding. We infer from these results that mRNA/gRNA binding is thermodynamically favorable and the U-tail contributes by forming and constraining new base pairs.

## **INTRODUCTION**

*Trypanosoma brucei* survival depends on developmentally regulated, post-transcriptional editing of mitochondrial mRNAs. The modification involves insertion and deletion of uridylates (Us) to generate functional open reading frames (Benne et al. 1986; Madison-Antenucci et al. 2002; Simpson et al. 2003; Stuart et al. 2005). Editing is catalyzed by a multiprotein complex, the editosome (Carnes et al. 2008) and requires a guide RNA (gRNA) to precisely direct the events. gRNAs are small mitochondrial-encoded

transcripts (50-70 nt) with complementary sequence (allowing G-U base pairing) to the edited message (Sturm and Simpson 1990). The enzyme cascade model for kinetoplastid mRNA editing starts with a site-specific endonuclease cleavage by the editosome after recognition of the mRNA/gRNA hybrid (Stuart et al. 2005). Thus, mRNA/gRNA complex formation is fundamental for the process. In *T. brucei*, the amount of RNA editing is substantial, involving 12 mRNA genes and at least 1200 gRNAs (Benne 1994; Stuart et al. 2005; Ochsenreiter et al. 2008). The pathways leading to complex formation for the hundreds of mRNA/gRNA pairs and assembly with the editosome are yet to be determined.

The absence of conserved sequences within mRNAs and gRNAs leads researchers to believe that RNA structure is important for proper complex formation and editing. Owing to mounting evidence that emphasizes the importance of RNA structure in a wide variety of biological processes, the secondary structure for individual and hybridized mRNAs and gRNAs has been receiving increasing attention (Piller et al. 1995; Schmid et al. 1995; Hermann et al. 1997; Connell and Simpson 1998; Leung and Koslowsky 1999; Cruz-Reyes et al. 2001; Leung and Koslowsky 2001a; 2001b; Pai et al. 2003; Simpson et al. 2003; Koslowsky et al. 2004; Cifuentes-Rojas et al. 2005; Yu and Koslowsky 2006; Cifuentes-Rojas et al. 2007; Zikova et al. 2008).

The primary structure of gRNAs comprises three domains, the 5' anchor (~4-21 nt), the central guiding region, and the 3' poly(U)-tail (10-20 nt). As a single molecule, the gRNAs are thought to fold into a small stem-loop (SL I) at the anchor, a second and larger stem-loop (SL II) within the guiding region, and a single stranded U-tail (Schmid et al. 1995; Hermann et al. 1997; Golden and Hajduk 2006; Yu and Koslowsky 2006).

Each mRNA editing domain has an anchor binding site (ABS) and a GC rich region but not much is known about its structure. The secondary structure for a few editing domains was predicted to form stem-loops involving the ABS and the initial editing sites (Piller et al. 1995; Connell et al. 1997; Koslowsky et al. 2004). The secondary structures for two truncated mRNAs, used as transcripts for in vitro editing experiments, the apocytochrome b (CYb) and ATPase subunit 6 (A6) have been solved and both had a stem-loop around the first editing site (Yu and Koslowsky 2006; Reifur and Koslowsky 2008). While there is indication of a cis-element within CYb and NADH dehydrogenase 7 (ND7) mRNAs in *Leishmania tarentolae* (Kabb et al. 2001; Oppegard et al. 2003) the mRNA secondary structure, not sequence, was found to be determinant of both gRNA binding affinity and editing efficiency in *T. brucei* (Koslowsky et al. 2004; Cifuentes-Rojas et al. 2005; Cifuentes-Rojas et al. 2007).

In vitro, A6 and CYb mRNA/gRNA pairs hybridize with distinct binding affinities (Koslowsky et al. 2004) and form a similar structure that is maintained after editing of the first three sites (Leung and Koslowsky 1999; 2001a; 2001b; Yu and Koslowsky 2006; Reifur and Koslowsky 2008). The editing-initiating gRNA for A6 is gA6-14, which has a strong affinity for A6, especially when compared to gCYb-558, the gRNA for CYb. Despite the difference in affinity, the U-tail was shown to augment the binding affinity for both A6 and CYb complexes and also for two other ND7 complexes (Koslowsky et al. 2004; chapter 3 of this thesis). The native A6/gA6-14 complex is used as a gold standard in editing assays in vitro because it is the only unmodified pair able to anneal and undergo a full round of editing. In vitro annealing of the A6 pair can be further improved in the presence of MRP (Muller et al. 2001; Zikova et al. 2008) or the

RBP16 (Ammerman et al. 2008) proteins. However, the A6 mRNA does not require these proteins to be edited in vivo, as determined by RNAi studies (Vondruskova et al. 2005; Zikova et al. 2006; Ammerman et al. 2008; Zikova et al. 2008). Based on the observations above, we infer that thermodynamics of mRNA/gRNA interaction is important within the editing context.

With purpose of studying the binding thermodynamics of A6/gA6-14, we utilized isothermal titration calorimetry (ITC). ITC determines the thermodynamic parameters of a bimolecular interaction in solution without the need of labeling or immobilizing any of the RNAs. Our results indicate that the mRNA/gRNA interaction is enthalpically favored, as expected, and the U-tail contributes by forming and constraining new base pairs.

## **RESULTS**

### **Description of the RNAs**

Amongst the several editing domains within the A6 transcript, we used the 3' most domain, which is the model transcript for editing assays in vitro. The A6 transcript was designed to be partially edited at site 1 (A6P1) to facilitate comparison with previously published experiments (Koslowsky et al. 2004) conducted with the same mRNA. Editing of this domain is initiated by gA6-14 (65 nt) that directs the editing of 81 U insertions and 8 deletions within 34 editing sites. Both RNAs (Fig. 4.1) have been previously described (Koslowsky et al. 2004) and extensively used in editing assays in vitro (Seiwert and Stuart 1994; Cruz-Reyes and Sollner-Webb 1996; Kable et al. 1996; Seiwert et al. 1996).

```

mRNA - A6P1 (67 nt)
5'GGAAAGGUUAGGGGGAGGAGAGAAGAAAGGGAAA
GUUGUGA**UUGGAGUUUAUAGAAUAAGAUCAAAUA3'

gRNA - gA6-14 (65 nt)
5'GGACUAUAACUCCGAUAACGAAUCAGAUUUUGAC
AGUGAUUAUGAUAAUUUUUUUUUUUUUUUUUUUU3'
sU: no U-tail
C6: 5'...UUUCCCUUUCUUCUC3'

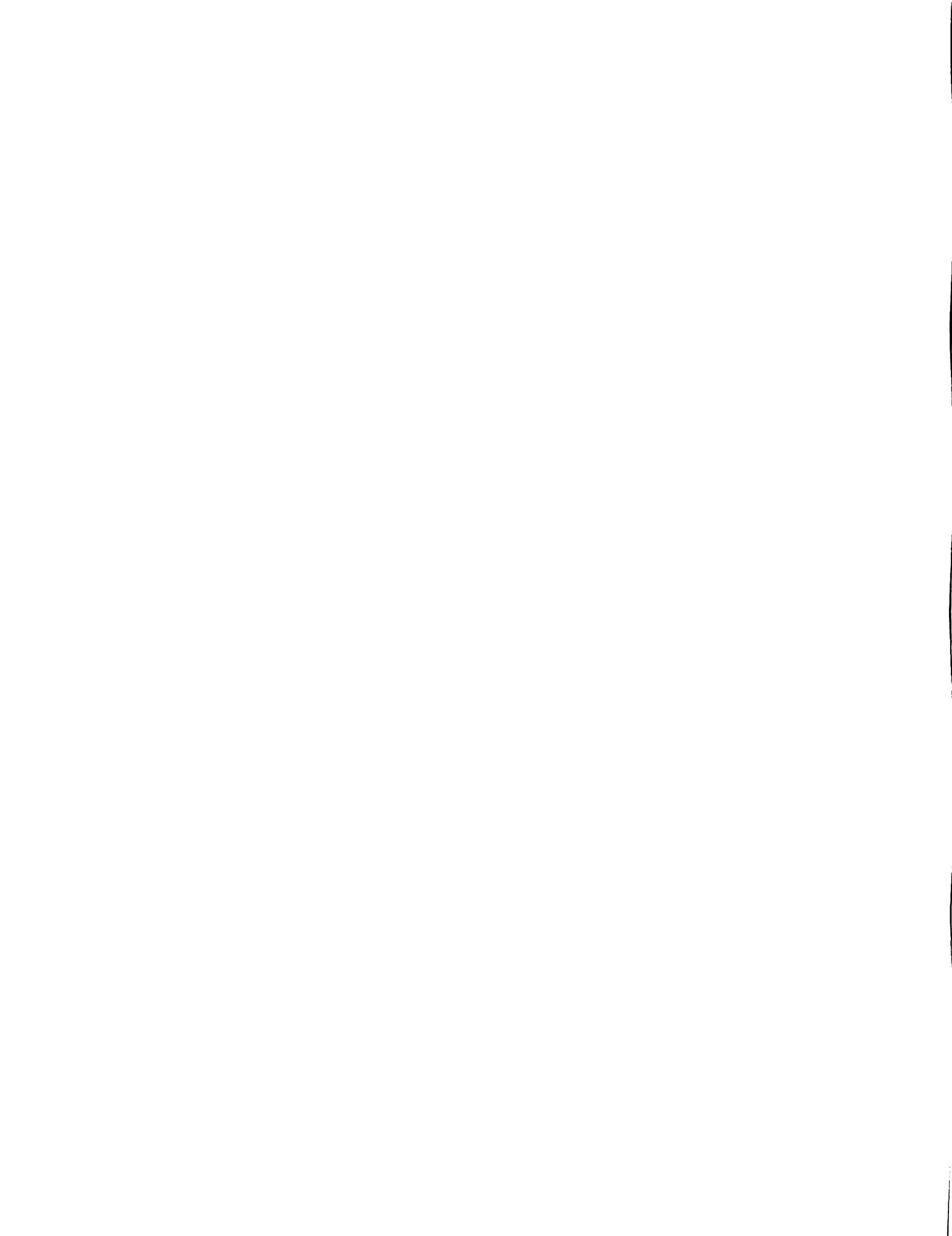
```

Figure 4.1 mRNA and gRNA sequences.

The mRNA anchor binding site (ABS) and the gRNA anchor are highlighted. A6P1 is partially edited at site 1 (asterisks indicate the two deleted Us). Underlined sequence: U-tail. Suffix sU: absence of U-tail, and C6: modified U-tail (6 U-to-C mutations).

### Individual RNA structures and Computer simulations

A6P1 (67 nt) is predicted to be mostly unstructured (Fig. 4.2A), with a single GU-rich 7 bp-stem-loop that includes the ABS (Koslowsky et al. 2004). The secondary structures for all of the gA6-14 constructs (Fig. 4.2A) have been confirmed by enzymatic and chemical probing (Reifur and Koslowsky 2008). gA6-14 (65 nt) forms two stem-loops, SL I and II. SL I is small (3 bp-stem) and located within the 5' anchor region. SL II is larger (7 bp-stem), and is formed by the guiding region of the molecule. In this gRNA, the U-tail (U15-tail) has 15 Us and is single stranded. gA6-14sU (50 nt) is a gA6-14 construct that lacks the U-tail. SL I and SL II are present in gA6-14sU but the molecule is less stable due to the absence of the U-tail. gA6-14-C6 (65 nt) is another construct that has a modified U-tail (C6-tail) containing 6 U-to-C mutations. The C6-tail promotes overall structural changes to a more stable conformation containing 3 stem-loops. Note in figure 4.2A that the C6-tail is not single stranded. This gRNA, previously named g[2,1]+6C, increases the editing efficiency of an in vitro reaction by a factor of 4 (Cruz-Reyes et al. 2001). We utilized mfold, DINAMelt (Zuker 2003; Dimitrov and Zuker 2004; Markham and Zuker 2005), and RNAstructure 4.6 (Mathews et al. 2004) to obtain



predicted free energies of the RNAs, to confirm the structures, and to obtain base pair probabilities.

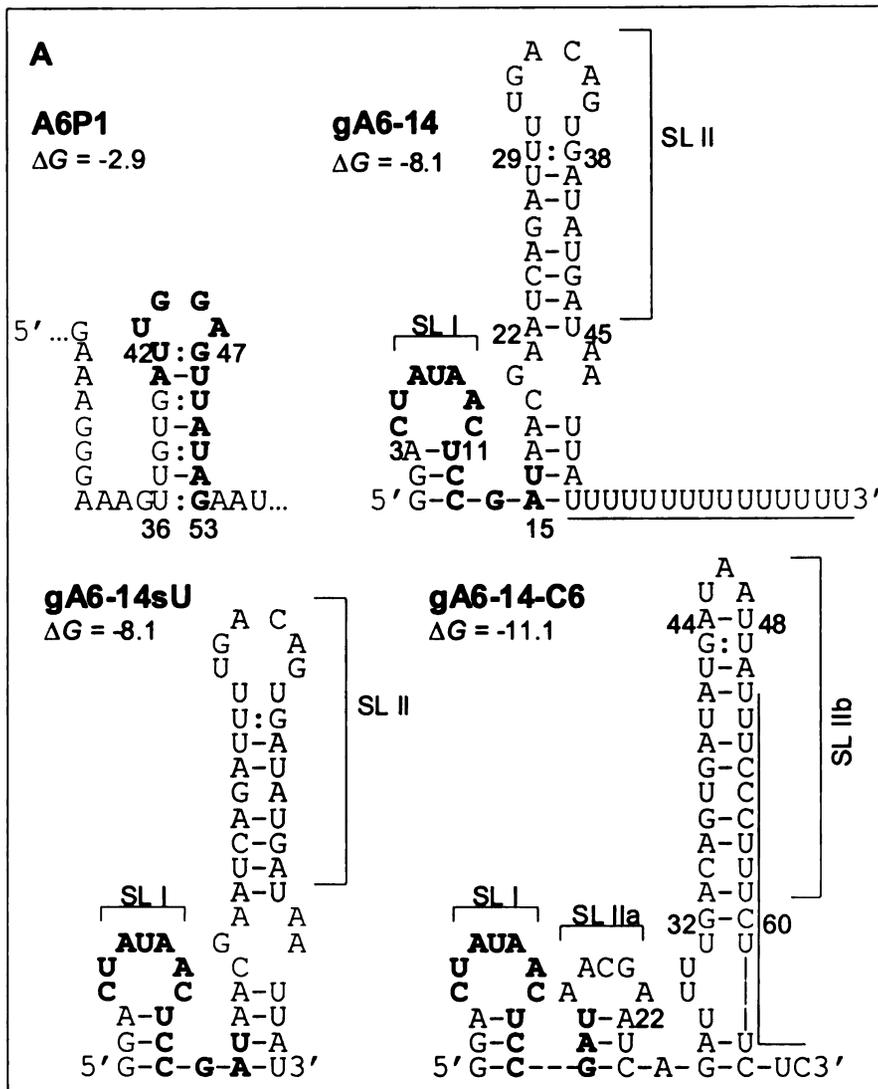


Figure 4.2 Secondary structures of the mRNA, gRNAs and complexes.

**A)** Individual RNAs: A6P1 and gRNAs (gA6-14, gA6-14sU and gA6-14-C6). The predicted changes in free energies ( $\Delta G$ , in Kcal/mol) were calculated by DINAMelt, at 27°C. A6P1 is predicted to be single stranded 5' and 3' of the indicated stem-loop. The highlighted regions within the A6P1 and gRNAs indicate the ABS and anchor, respectively. The U-tail and C6-tail are underlined and the conserved stem-loop (SL II) is shaded. SL IIa and SL IIb are only present in gA6-14-C6. **B)** Bimolecular mRNA/gRNAs. Shaded boxes in gA6-14sU and gA6-14-C6 represent matches with gA6-14. (\*) ES. (-) Watson-Crick bonds. (: ) wobble pairing.



## **mRNA/gRNA structure**

Work from our laboratory confirms formation of a three-helical structure at 27°C and 10 mM Mg<sup>+2</sup> for the unedited A6U/gA6-14 complex in solution (Reifur and Koslowsky 2008). The structure remained unaltered when gA6-14 was substituted by gA6-14-C6 or gA6-14sU, with the exception that the U-tail helix could not form with gA6-14sU due to the absence of the U-tail. Whereas the secondary structure is less stable upon deletion of the U-tail, substitution with the C6-tail promotes higher stability. A6P1 is identical in sequence to A6U except for the deletion of two uridylates at ES1. Computer simulations indicate that the uridylate deletions do not substantially affect the structure of the mRNA/gRNA interactions. A6P1 hybridization with gA6-14 is predicted to form a three-helical secondary structure similar to that observed for A6U. The 3'-end of A6P1 contains the anchor binding site (ABS), which is complementary to the gRNA anchor. Both strands are predicted to hybridize with a probability of 100%, calculated by DINAMelt. This annealing forms the anchor helix, which is sufficient for editing initiation (Seiwert et al. 1996). Just upstream of the anchor helix is the first editing site (ES) for A6P1, identified as the first mismatch 5' of the anchor duplex. The probability of the base at the ES (G<sub>40</sub>) to be involved in a base pair is zero. Upstream of the ES, 4 bp are predicted to form between the guiding nucleotides and the mRNA with a probability of 95-97%. Approximately 6 nt upstream of ES is the sequence that most frequently interacts with the gRNA U-tail and forms the 5' U-tail helix. SL II within the gRNA contributes to the formation of the third helix at the junction. Although figure 4.2B depicts the U-tail bound to a specific region within A6P1/gA6-14, this interaction is known to be flexible (Reifur and Koslowsky 2008). A dynamic U-tail interaction with

the mRNA has also been shown for CYb unedited and partially edited complexes (Leung and Koslowsky 1999; Yu and Koslowsky 2006). To counteract the flexibility of the U-tail helix, we utilized gA6-14-C6 (Cruz-Reyes et al. 2001). The six U to C mutations lock the U-tail into a single position (Fig. 4.2B). Predictions using DINAMelt with default parameters confirm that the U-tail promotes stability to the complex (Table 4.1).

Table 4.1 Predicted changes in free energy ( $\Delta G$ ) and melting temperature ( $T_m$ ) for the A6 mRNA/gRNA complexes used in this study.\*

mRNA/gRNA	$\Delta G^\circ$ (Kcal/mol)*
A6P1/gA6-14	-35.59
A6P1/gA6-14sU	-30.55
A6P1/gA6-14-C6	-46.94

\*values are for the ensemble, which includes contributions from single stranded species and homo-dimers. It does not include intramolecular pairings or tertiary interactions when the RNA strands are base paired. However, it does include intramolecular pairing within individual RNAs.

## Determination of binding affinity

The apparent binding affinities ( $K_D$ ) for A6P1 with gA6-14 and gA6-14sU have been previously measured using electrophoretic mobility shift assays (EMSA) (Koslowsky et al. 2004). In these experiments, the A6P1/gA6-14 interaction was studied at various concentrations of magnesium using a total incubation time of 45 minutes. Under these conditions, the relative binding affinity of gA6-14 for its mRNA was high, with an observed  $K_D$  of approximately 5 nM. Neither the  $Mg^{+2}$  concentration nor the presence/absence of a U15 tail had a major affect on the interaction, with a single predominant band forming under all conditions tested (Koslowsky et al. 2004).

Subsequent analyses indicated that under low magnesium conditions, 45 minutes may not be long enough for the interaction to reach equilibrium. Therefore, in order to accurately determine the thermodynamic contribution of the U-tail to the bimolecular interaction at

2 mM  $Mg^{+2}$  (ITC conditions) the EMSA experiments were repeated with the three gA6-14 constructs described above. Briefly, each pair was combined (5 nM of  $^{32}P$ -labeled gRNA and increasing concentrations of mRNA), denatured at 70°C for 2 min, slow cooled to 27°C and allowed to anneal for 3 hours in a buffer containing 2 mM  $Mg^{+2}$ . The samples were then loaded under current onto nondenaturing 6% polyacrylamide gels. The gel images clearly show the shift of a single major complex for all three mRNA/gRNA pairs (Fig. 4.3). In addition, multiple slow-migrating minor complexes were visible. Diffuse signal between the main complex and the free gRNA indicates that dissociation does occur during electrophoresis. The main complexes were quantified and the apparent dissociation equilibrium constant ( $K_D$ ) calculated. The calculated  $K_D$  for gA6-14 (with the U15 tail) was 2.7 nM, approximately half that determined previously, indicating that at 45 minutes the 2 mM  $Mg^{+2}$  reaction was in fact not at equilibrium. In the absence of the U-tail (gA6-14sU) the observed  $K_D$  was approximately 2-fold higher (5.8 nM) indicating that the U-tail contributes significantly to the RNA interaction. As expected, addition of the C6-tail significantly stabilized the interaction, dropping the observed  $K_D$  almost 10-fold (0.2 nM).

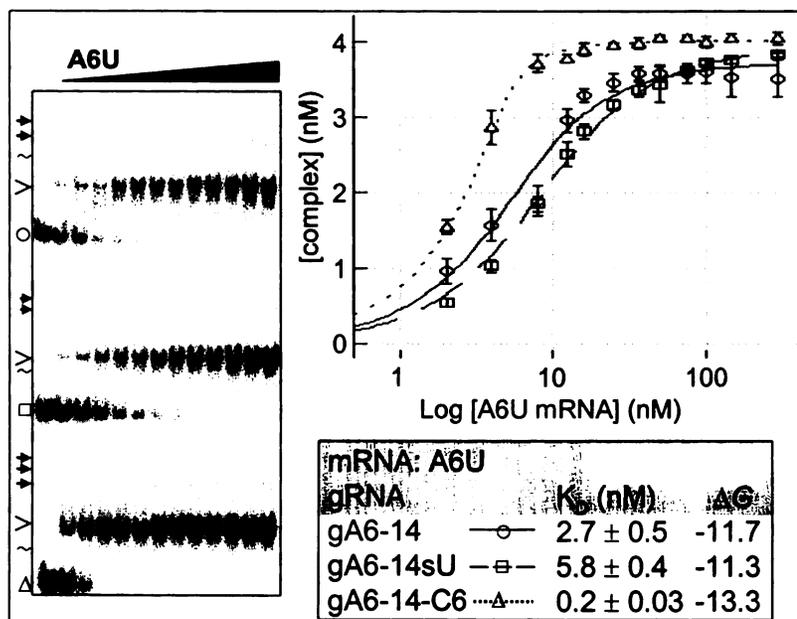


Figure 4.3 mRNA/gRNA complex formation by EMSA.

Samples contained 5 nM  $^{32}\text{P}$ -labeled gRNA and increasing concentrations of the A6U mRNA (from left to right: 0, 2 nM, 4 nM, 8 nM, 12.5 nM, 16 nM, 25 nM, 37 nM, 50 nM, 75 nM, 100 nM, 150 nM, and 300 nM). Free gA6-14 ( $\circ$ ), gA6-14sU ( $\square$ ), gA6-14-C6 ( $\Delta$ ) and minor complexes of slow mobility, dependent ( $\rightarrow$ ), or independent ( $\sim$ ) of 6U were observed at the indicated positions. Quantification of main complex formation ( $>$ ) allowed calculation of the apparent  $K_D$ . The binding curve shows the average result obtained from 4 experiments for each pair of mRNA/gRNA. Error bars indicate the standard deviation in complex formation.  $\Delta G$  is in Kcal/mol, as estimated by  $-\text{R T} \ln (K_D^{-1})$ .

## Thermodynamics of mRNA/gRNA binding

ITC allows the measurement of change in enthalpy ( $\Delta H$ ), change in entropy ( $\Delta S$ ), association constant ( $K_A$ ), and the stoichiometry of the reaction ( $n$ ), by measuring stepwise changes in enthalpy during the course of a titration experiment at a constant temperature. From these parameters, one can derive the change in free energy ( $\Delta G$ ) and the dissociation equilibrium constant ( $K_D$ ) by using the equations  $\Delta G = \Delta H - (T \Delta S)$ , where T is the Kelvin temperature, and  $K_D = K_A^{-1}$  (Wiseman et al. 1989).

ITC was used here to measure the thermodynamics of mRNA/gRNA complex formation under physiologically relevant salt conditions (2 mM  $Mg^{+2}$ ) and temperature (27°C). We chose the A6 mRNA/gA6-14 pair because they are known to anneal into an active complex in RNA editing assays in vitro without auxiliary factors (Seiwert and Stuart 1994; Kable et al. 1996; Seiwert et al. 1996; Koslowsky et al. 2004). Upon gA6-14/A6P1 annealing, we predicted the breakage of a few intramolecular hydrogen bonds within the mRNA and gRNA, and formation of a higher number of new intermolecular bonds to create the anchor and the U-tail helices shown in figure 4.2. Formation of hydrogen bonds releases heat to the surrounding (free energy increments of -0.5 to -1.5 Kcal/mol per hydrogen bond), which makes it an exothermic interaction that promotes RNA stability (Turner 1999). Therefore, we expected complex formation to be driven by a favorable negative  $\Delta H$ . These bonds would also stabilize the conformation by reducing phosphate backbone rotations of the complex, decreasing the mRNA free-energy and entropy. Using gA6-14sU, only the anchor helix and gRNA stem-loop are present and the overall complex is less stable, thus we predicted to obtain  $\Delta H$ ,  $\Delta G$  and  $\Delta S$  less negative than gA6-14. gA6-14-C6 on other hand, provides formation of more hydrogen bonds within the U-tail helix (6 C-G Watson-Crick vs ~6 U:G Wobble bonds), providing stronger conformational stabilization and immobilization of the complex.

In initial ITC experiments we utilized mRNA and gRNA concentrations based on EMSA's measured  $K_D$  to obtain a binding isotherm with the  $c$  value within the optimal experimental window of  $5 \leq c \leq 500$  (according to VP-ITC MicroCalorimeter User's Manual and Wiseman et al 1989).  $c$  is the product of the association constant ( $K_A$ ) and the concentration of the receptor ( $[M]$ ) when the receptor has one binding site ( $n = 1$ ) or  $c$

$= n K_A [M]$ . From the  $K_D$  measured by EMSA (5 nM) we determined the  $K_A$  (0.2 nM) using the equation  $K_A = K_D^{-1}$ . Aiming for a  $c$  value of 25, we used the mRNA (receptor) in the sample cell at a concentration of 125 nM and the gRNA in the syringe at a concentration 10 times higher than the mRNA. Surprisingly, the binding event under these concentrations (mRNA = ~125 nM and gRNA = ~12.5  $\mu$ M) did not evolve more heat than the background noise (data not shown). This was an indication that the  $K_D$  measured by EMSA was lower than what the ITC was measuring or the  $\Delta H$  was too small to be detected at these concentrations. With increased concentrations however, we were able to obtain binding isotherms with  $c$  values within the optimal experimental window.

### **A6P1/gA6-14sU interaction**

Titration of gA6-14sU into the sample cell containing A6P1 (three identical experiments) resulted in heat evolved (negative peaks), confirming exothermic binding. Integration of the heat measurements (corrected for dilution of gA6-14sU into buffer) generated a steep titration curve with binding saturation obtained after the ~12<sup>th</sup> injection (Fig. 4.4).

Titration of gA6-14sU into A6P1 should have produced a 1:1 complex; however, the calculated value of  $n$  was smaller than 1 (~ 0.6), indicating depletion of mRNA. Fitting the data to a single binding site model (Holdgate 2001) generated a good curve fit and resulted in a  $\Delta H$  of  $-47.3 \pm 2.6$  Kcal/mol and a  $\Delta S$  of  $-123 \pm 8.6$  eu (Fig. 4.4A, Table 4.2).

The  $\Delta G$  of  $-10.9 \pm 0.9$  Kcal/mol was closer to the  $\Delta G$  calculated by EMSA (-11.3 Kcal/mol). The apparent  $K_D$ s calculated from the different runs were quite variable, ranging from 24.3 nM to 57.5 nM. However, in these analyses, we focused on the fitted  $\Delta H$ , as in high affinity interactions (i.e.  $K_A \geq 10^8$ ), the  $K_A$  value is determined from the

binding of the last ~5% of strands in the sample cell and cannot be confidently fit by this method (Wiseman et al. 1989; Mikulecky et al. 2004).

### **A6P1/gA6-14 and A6P1/gA6-14-C6 interactions**

Titration of both gA6-14 and gA6-14-C6 into the sample cell with A6P1 also resulted in an exothermic binding that reached saturation at a gRNA to mRNA molar ratio smaller than 1. In both cases, the  $\Delta S$  and  $\Delta H$  were more negative than gA6-14sU, as expected (Fig. 4.4A, Table 4.2). However, for both gRNA constructs, integration of the heat measurements generated a titration curve with a suboptimal fit to the single binding site model.

The suboptimal fit was due to a distortion of the curve observed in all ITC experiments using the gRNAs that contained a tail, and hence two mRNA interaction sites (U15-tail or the C6-tail). The distortion caused by the gRNAs with a tail resembled what is seen for an additional binding mode (Chaires 2001), as if the mRNA or the gRNA had two binding sites. Both integrated curves could be fit to the two binding site model using the Origin software package (Fig. 4.4B, Table 4.2). Surprisingly, the curve fits indicated that it was the “lower” affinity site that titrated in the early injections. This suggests that introduction of a second mRNA binding site on the gRNA introduced unexpected basepairing interactions or artifacts. Upon injection, the ITC measures the heat evolved or taken up due to basepair breakage (endothermic), basepair formation (exothermic), as well as heats of dilution and mixing. We speculate that in the early injections (very high mRNA to gRNA ratios), the introduction of a second binding site on the gRNA may allow for the interaction of two mRNAs with a single gRNA. Movement of this possible intermediate to a resolved mRNA/gRNA complex would involve

breakage of hydrogen bonds, hence in the initial injections less total heat is evolved than expected. The thermodynamic parameters determined from the two site model for the second, high affinity binding site, show considerable variation between experiments (data not shown); an indication that the two site model is incorrect. In order to more accurately determine the thermodynamic contribution of the U-tail, the data was re-analyzed using the single binding site model after dropping the initial (low affinity) data points (Fig. 4.4B, Table 4.2). For gA6-14, these analyses generated much more consistent numbers with  $n$  values closer to 1. The calculated  $\Delta H$  for gA6-14 was  $-64.1 \pm 5.6$  Kcal/mol, indicating that the U-tail contributes with a significant enthalpic component. A similar analysis of the C6 gRNA construct indicated that the U to C substitutions further increased the enthalpic contribution of the U-tail as expected. However, there was still variation between experiments. The calculated  $\Delta H$  for three different runs varied from -74.7 Kcal/mol to -112 Kcal/mol but the calculated  $n$  values were higher.

Table 4.2 Thermodynamic parameters from isothermal titration calorimetry.

<b>ONE SITE BINDING MODEL</b>								
gRNA	mRNA [ $\mu$ M]	gRNA [ $\mu$ M]	$-\Delta H$ cal/mol	$-\Delta S$ eu*	$-\Delta G$ Kcal/mol	$K_A$ $M^{-1}$	$K_D$ nM	$n$
sU	3.2	32.4	4.46E4 $\pm$ 663	114	12	3.06E7	57.5	0.7
	3.8	43.4	4.77E4 $\pm$ 635	125	10.2	2.79E7	35.8	0.7
	4.4	36.1	4.98E4 $\pm$ 514	131	10.5	4.12E7	24.3	0.5
mean			4.73E4	123	10.9	3.3E7	39.2	0.6
$\pm$ SD			$\pm$ 2E3	$\pm$ 8.6	$\pm$ 0.9	$\pm$ 7E6	$\pm$ 16.8	$\pm$ 0.1
14	3.1	26.4	8.45E4 $\pm$ 2.5E3	250	9.9	1.82E7	54.9	0.6
	3.1	35.5	7.65E4 $\pm$ 1.8E3	222	9.9	1.77E4	56.5	0.7
	3.3	27.5	7.31E4 $\pm$ 2.1E3	211	9.8	1.64E7	61.0	0.6
mean			7.71E4	224.4	9.9	18.0 $\pm$	55.5	0.6
$\pm$ SD			$\pm$ 9.3E3	$\pm$ 31.0	$\pm$ 0.1	0.3	$\pm$ 3.5	$\pm$ 0.04
C6	2.5	18.1	1.71E5 $\pm$ 1.0E4	539	9.1	4.68E6	213.7	0.5
	2.7	18.4	1.61E5 $\pm$ 4.9E3	505	9.6	1.25E7	80.0	0.6
	4.4	32.9	1.39E5 $\pm$ 7.4E3	432	9.3	6.50E6	153.8	0.4
mMan			1.72E5	542	9.2	6.8E6	181	0.45
$\pm$ SD			$\pm$ 3.2E5	$\pm$ 109	$\pm$ 0.2	$\pm$ 3.9E6	$\pm$ 84.5	$\pm$ 0.1
<b>ONE SITE BINDING MODEL AFTER DELETION OF INITIAL TITRATIONS</b>								
gRNA	mRNA [ $\mu$ M]	gRNA [ $\mu$ M]	$-\Delta H$ cal/mol	$-\Delta S$ eu*	$-\Delta G$ Kcal/mol	$K_A$ $M^{-1}$	$K_D$ nM	$n$
14	3.1	26.4	6.09E4 $\pm$ 1E4	170	9.9	1.85E7	57.4	0.91
	3.1	35.5	7.06E4 $\pm$ 2.7E3	202	10	2.37E7	42.2	0.79
	3.3	27.5	6.07E4 $\pm$ 3.1E3	168	10.3	2.96E7	33.8	0.73
mean			6.41E4	180	10.07	2.39E7	44.47	0.81
			$\pm$ 5.6E3	$\pm$ 19	$\pm$ 0.21			$\pm$ 0.09
C6	2.5	18.1	8.06E4 $\pm$ 2.5E3	234	10.4	3.73E7	26.8	0.76
	2.7	18.4	1.12E5 $\pm$ 5.6E3	339	10.3	3.78E7	26.5	0.68
	4.4	32.9	7.47E4 $\pm$ 4.2E3	214	10.5	3.45E7	29.0	0.56
mean			8.91E4	262.3	10.4	3.65E7	27.43	0.67
			$\pm$ 2E4	$\pm$ 67	$\pm$ 0.1			$\pm$ 0.1
Titrant: gA6-14sU, gA6-14, gA6-14-C6. In sample cell: A6P1. *eu = entropy unit = cal mol <sup>-1</sup> K <sup>-1</sup>								

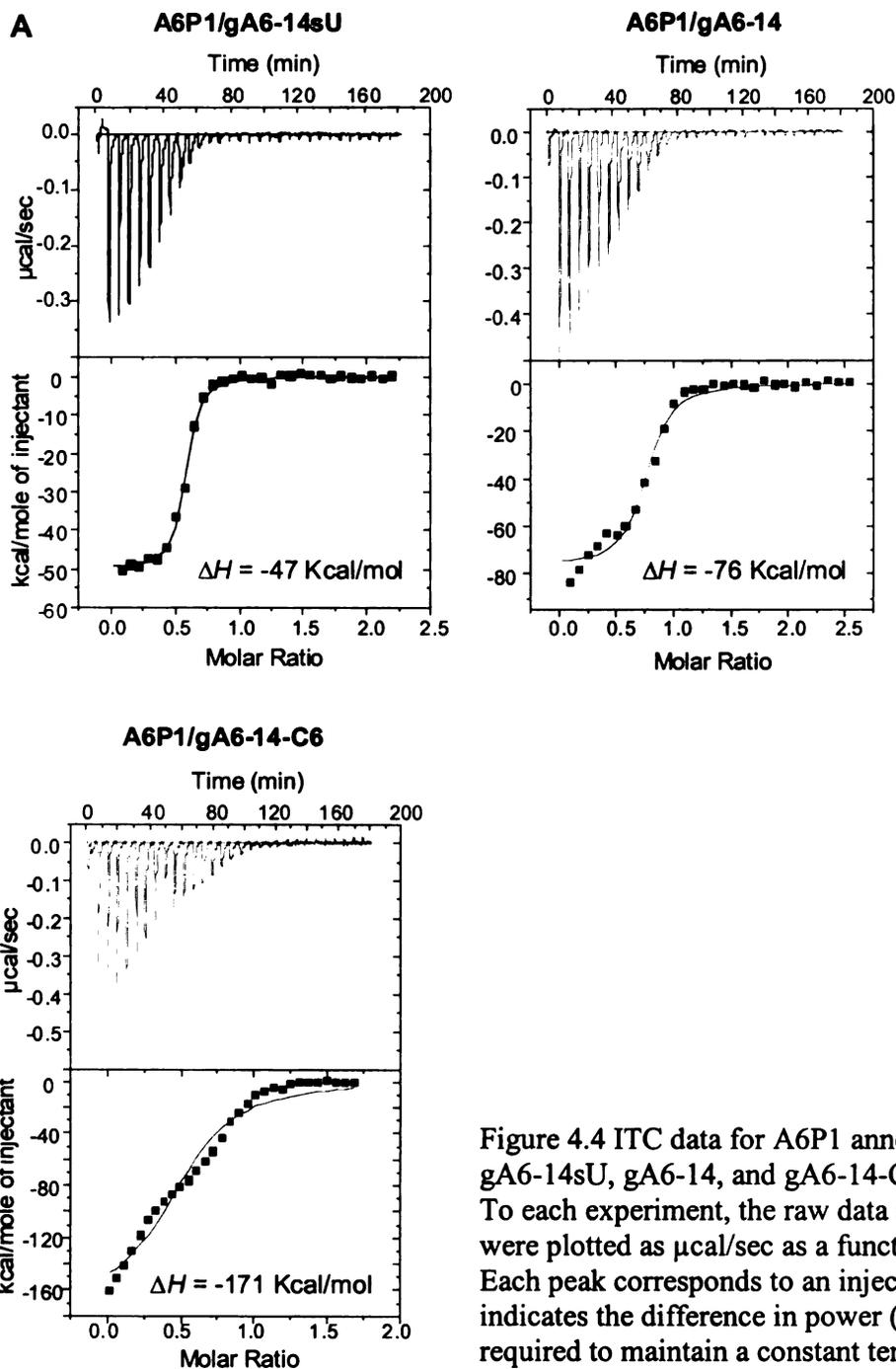


Figure 4.4 ITC data for A6P1 annealing with gA6-14sU, gA6-14, and gA6-14-C6. To each experiment, the raw data (upper peaks) were plotted as  $\mu\text{cal}/\text{sec}$  as a function of time. Each peak corresponds to an injection and indicates the difference in power ( $\mu\text{cal}/\text{sec}$ ) required to maintain a constant temperature between the sample cell and the reference cell.

The integrated data (lower panel) includes the raw data integrated to yield injections enthalpies (Kcal/mol), normalized for the moles of titrant gRNA added (gRNA to mRNA ratio). Squares indicate normalized injection data as a function of molar ratio. **A.** Data points fit (solid line) to a one-site binding model. **B.** Due to suboptimal line fit for gA6-14 and gA6-14-C6, data were also analyzed using the two-site model and the one-site model after deletion of initial titrations.

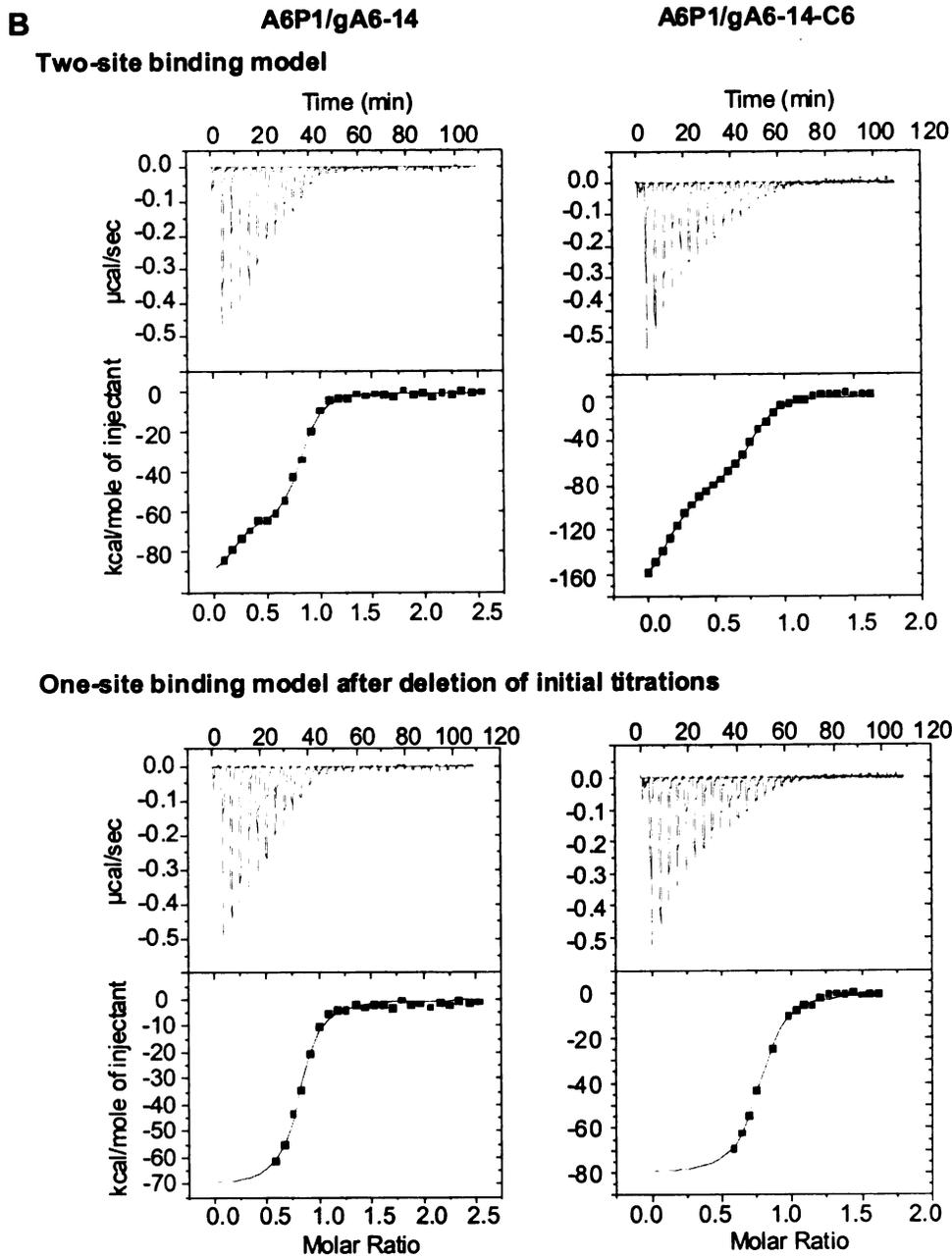


Figure 4.4 (cont'd).

The variations observed among experiments with the same RNAs and the  $n$  values smaller than 1, led us speculate potential experimental problems. The most common problems reported in ITC experiments with RNA are inaccuracy in calculating sample concentration and mismatched buffers. All RNAs utilized were extensively dialyzed, and the heat measured in the last few titrations indicated good buffer match, generating good baselines. The small  $n$  values could be suggestive of error in RNA sample concentration or degradation (Feig 2007). Radioactive labeling of input RNA after ITC measurements (data not shown) indicated that >90% of our RNAs were intact (both the mRNA in the sample cell and all injected gRNAs were analyzed) indicating that sample degradation could not explain the low  $n$ -values. Gel shift analyses of A6P1 with gA6-14, gA6-14sU, and gA6-14-C6, at 2 mM  $Mg^{+2}$ , resulted in the shift of a single predominant band (Koslowsky et al. 2004). However, as the mRNA concentration increased from 1.25 nM to 50 nM (constant gRNA concentration of 5 nM) two other less abundant bands of slower mobility appeared in the same lane. The second band was not analyzed for molecularity, thus it could involve trimolecular interactions, homo-dimers, or the bimolecular mRNA/gRNA adopting a conformation of slower mobility. Because the ITC experiments were conducted at very high RNA concentrations, we decided to repeat the gel shift experiments using the RNA concentrations as in ITC to clarify the above potential problems. In these experiments, no unusual complexes at high mRNA to low gRNA ratios were resolved (data not shown).

## **DISCUSSION**

While the structure and binding affinity of different mRNA/gRNA pairs have been reported, very little is known about the thermodynamic forces that drive the interaction.

The binding affinities of the A6P1/gA6-14 pair, measured by electrophoretic mobility shift assays at 2 mM  $Mg^{+2}$  indicate that the gRNA U-tail does contribute to the overall stability of the interaction (apparent  $K_{DS}$  of 2.7 nM for gA6-14 and 5.8 nM for gA6-14sU). In similar EMSA studies with other gRNA/mRNA pairs, the U-tail promotes a much more pronounced effect under physiological  $Mg^{+2}$  conditions (Koslowsky et al. 2004). For instance, the binding affinity of gCYb-558 for its cognate CYb mRNA ( $K_D$  with U-tail = 600 nM) more than doubled when the U-tail was absent ( $K_D$  without the U-tail = 1.3  $\mu$ M). Analysis of the rate constants using surface plasmon resonance (SPR) indicated that the weak binding affinity observed for the CYb/gCYb-558 pairs was because gCYb-558 bound very slowly to the highly structured CYb mRNA, producing an association rate constant ( $k_{on}$ ) of  $5.7 (\pm 0.9) \times 10^2 M^{-1} s^{-1}$ , and dissociation ( $k_{off}$ ) of  $2.7 \times 10^{-3} s^{-1}$  (Yu 2006). In this same study, the relatively open A6 mRNA triggered a much faster  $k_{on}$  [or normal when compared to other RNA-RNA interactions (Diamond et al. 2001)] of  $1.2 (\pm 0.2) \times 10^4 M^{-1} s^{-1}$  and  $k_{off} = 3.0 (\pm 1.5) \times 10^{-5} s^{-1}$ , which contributed to the lower  $K_D$ .

In this study, we directly measured the enthalpic contribution of the U-tail to the RNA interaction for gA6-14 using isothermal titration calorimetry. Our results indicate that the U-tails provides a significant enthalpic contribution to the interaction.

The ITC experiments using gA6-14sU generated steep binding curves with very good fits to a one-site binding model. In contrast, experiments with gA6-14 (with U15-tail) always generated titration curves with a slight inflection at ~0.5 molar ratio of gRNA to mRNA and a sub-optimal fit to the one-site model. Initially, we hypothesized that the flexibility of the U-tail interaction may be introducing interaction artifacts. However,

utilization of the C6 gRNA construct in order to define the U-tail position only intensified the inflection. This indicates that the introduction of a second interaction site on the gRNA caused the artifact in the titration curve, with the C6 gRNA construct causing a larger bend, likely due to the increased base-pairing abilities of the U to C modified U-tail. Analyses using a two-site binding model indicated that the early titrations involved a lower affinity interaction indicating that less heat was evolved than expected. We speculate that the second binding site on the gRNA can introduce RNA interaction artifacts caused by the very high concentrations of RNAs used during the ITC experiments. Resolution of the interaction artifacts involves additional endothermic hydrogen bond breakages, which would decrease the total enthalpic change. Re-analyzing the ITC data using the one-site model, but dropping the initial data points, did give us a good curve fit for both gA6-14 and gA6-14-C6. However, the gA6-14-C6 data still showed considerable amount of variation and the n-values were lower than those seen for gA6-14. The gA6-14-C6 gRNA differs substantially from gA6-14, in that the introduced U to C mutations change the secondary structure of the gRNA (Fig. 4.2A). The gA6-14-C6 structure is more stable, with a predicted  $\Delta G^\circ$  of -15.79 Kcal/mol, vs -11.72 Kcal/mol for gA6-14. This increase in stability and base pairs changes the pathway to mRNA/gRNA formation. The pathway includes an energy cost to partially open the gRNA and the mRNA, an energy cost for nucleation, and an energy gain from initial duplex formation and elongation. ITC measures the total enthalpy change of the binding process and this allows further calculation of total free-energy change where all possible interactions, including disruption and creation of secondary and tertiary interactions, are superposed. In addition, ITC also measures all the energy costs and gains from potential

alternative conformers that may interfere in the reaction. This may partially explain the variation in measured  $\Delta H^\circ$ . Small RNAs do form misfolded conformations, at least transiently, and states can alternate between active and inactive conformation with distinct kinetics (Zhuang et al. 2002; Russell 2008). It may be that the U to C mutations can kinetically trap the C6 gRNA decreasing the number of binding competent molecules.

This study provides for the first time the thermodynamic parameters for mRNA/gRNA binding event. The results definitively show that the U-tail can contribute with a significant enthalpic component helping to drive complex formation.

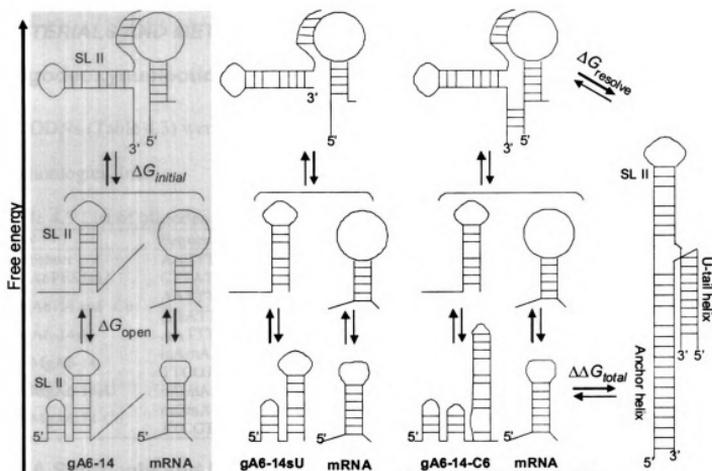


Figure 4.5 Schematics of structural rearrangements predicted to happen in the thermodynamics of three distinct mRNA-gRNA binding events. From left to right are the interactions of A6P1 mRNA with gA6-14, gA6-14sU and gA6-14-C6, respectively. The first step of binding involves partial mRNA and gRNA opening (unfavorable) to allow initial mRNA/gRNA contacts, leading to a more stable and resolved mRNA/gRNA complex (with gA6-14sU, the U-tail helix is not formed). The final free-energy change ( $\Delta\Delta G_{\text{total}}$ ) is the contribution of  $\Delta G_{\text{open}}$  (necessary to expose the binding site in the appropriate conformation),  $\Delta G_{\text{initial}}$  (initiation-energy penalty), and  $\Delta G_{\text{resolve}}$  (energy gain because of hybridization and tertiary interactions). ITC calculates  $\Delta\Delta G_{\text{total}}$ , which is the sum of  $\Delta G_{\text{open}}$ ,  $\Delta G_{\text{initial}}$ , and  $\Delta G_{\text{resolve}}$ . Not depicted in the chart are potential kinetically trapped alternative conformers and homo-dimers for the ssRNAs, and alternative conformers and higher-order-molecular interactions for the mRNA/gRNA complex, which contribute to the low stoichiometry measured by ITC.

## **MATERIALS AND METHODS**

### **Oligodeoxynucleotides (ODN)**

All ODNs (Table 4.3) were chemically synthesized and obtained from Integrated DNA Technologies, Inc.

Table 4.3 List of oligodeoxynucleotides

<b>ODN</b>	<b>Sequence (5' to 3')</b>
T7A6short	AATTTAATACGACTCACTATAGGAAAGG
Rev A6PES1sh2	CTTATTTGATCTTATTCTATAACTCCAATCAC
T7-gA6-14 and -C6	AATTTAATACGACTCACTATAGGACTATAACTCCGATAACGAATC AGATT
T7-gA6-14sU	AATTTAATACGACTCACTATAGGACTATAACTCCGATAACG
2-O-MgA6-14	mAmAAAAAAAAAAAAAAAAATAATTATCATATCACTGTCAAATCTGA TTCGTTATCG
2-O-MgA6-14sU	mUmAATTATCATATCACTGTCAAATCTGATTTCGTTATCGGAG
2-O-MgA6-14-C6	mGmAGAAGAAAGGGAAATAATTATCATATCACTGTCAAATCTGA TTCGTTATCG

### **RNA Synthesis and Labeling**

DNA templates for mRNA transcription were PCR amplifications (using T7 A6 short and Rev A6PES1sh2 primers) from plasmids described previously (Koslowsky et al. 2004). Template for gRNA transcription was obtained by hybridizing two overlapping ODNs (appropriate T7- and 2-O-M-) and adding Klenow fragment of DNA polymerase I (New England Biolabs, Inc.) plus dNTPs to make dsDNA from ssDNA templates, as per manufacturer's recommendation. The gRNA template was further amplified using PCR and the same ODNs as primers. The Ribose C2' methoxy groups (–OCH<sub>3</sub> or m) on the two last nucleotides of the 5' termini of the 2-O-M- ODNs were used to increase transcript homogeneity by inhibiting the T7 RNA Polymerase from adding one or more nontemplated nucleotides at the 3' end of the nascent transcripts (Pleiss et al. 1998; Helm et al. 1999; Kao et al. 2001). The RNAs were synthesized using the above DNA templates and T7 RNA polymerase (RiboMAX™ Large Scale RNA Production System,

Promega Co.) according to the manufacturer's directions. RNAs were gel purified on a 20% denaturing PAGE and 5'-end-labeled if necessary with T4 Polynucleotide Kinase (Invitrogen) and [ $\gamma$ <sup>32</sup>P] ATP. RNAs were suspended in ultra-pure water (MilliQ, Millipore) gel shift or ITC buffer. RNA concentration was determined at room temperature by measuring the UV absorbance at 260 nm, using a ND-1000 Spectrophotometer (NanoDrop Technologies, Wilmington, DE) and the equation  $A = \epsilon Cl$  ( $A$  = absorbance,  $C$  = concentration,  $\epsilon$  = extinction coefficient,  $l$  = UV path length). The extinction coefficient for each RNA was obtained online at <http://www.dharmacon.com/>.

### **Secondary Structure Prediction**

Predicted secondary structures and free energies were obtained using mfold, at the URL <http://mfold.bioinfo.rpi.edu/>, DINAMelt, at the URL <http://dinamelt.bioinfo.rpi.edu/>, (Zuker 2003; Markham and Zuker 2005), and RNAstructure 4.6, at the URL <http://rna.urmc.rochester.edu/rnastructure.html> (Mathews et al. 2004).

### **Isothermal Titration Calorimetry (ITC)**

Prior to an experimental run, all RNAs were dialyzed against the same buffer on Slide-A-Lyzer® cassettes, 0.5-3 mL capacity (Pierce). Dialysis was conducted stirring at 4°C on 600 mL of ITC buffer, including two buffer changes during the 6-8 h dialysis process. ITC hybridizing buffer: 2 mM MgCl<sub>2</sub>, 20 mM Tris pH 7.5, 100 mM KCl, and 0.1 mM EDTA. After dialysis, RNA concentration was determined as above. Before each experimental run, the ITC compartments and other accessory parts were thoroughly cleaned to remove any contaminants, specifically RNase. The sample cell was soaked with a solution of 5% Contrad-70 (Fisher Scientific) at 27°C for 10 min. The auto-

pipette was connected to the sample cell and rinsed with 400 mL of same Contrad solution at room temperature and then rinsed with 1 L of ultrapure water and left to dry and calibrate over night at 27°C. If after this procedure the results were not satisfactory, an additional step filling the cell and auto-pipette with RNase Zap (Ambion) followed by the water rinse. All the other equipment that had contact with RNA, such as the cell, syringes, tubings, vials, and spin bars were soaked in RNase Zap, rinsed five times with ultrapure water, and then left to dry over night at room temperature. After dialysis and data collection, RNA samples were assayed by denaturing (1.6 mM-thick) PAGE. To better assess sample degradation, we dephosphorylated the RNAs (using CIP phosphatase from NEB), then 5'-end labeled with [ $\gamma$ <sup>32</sup>P] ATP (as described above), and analyzed the radioactive samples by denaturing PAGE. ITC experiments were performed at 27°C on a VP-ITC titration calorimeter (MicroCal, Inc., Northampton, MA). The RNAs were heat denatured at 70-80°C, for 2 minutes, and kept on ice until they were degassed. Then, the mRNA (1.4 mL at 1.4-4.4  $\mu$ M) was transferred to the sample cell and the gRNA (~300  $\mu$ L at a concentration ~10 times that of the mRNA) was loaded into the auto-pipette. Each gRNA titration consisted of around 30 injections of 10  $\mu$ L each (first injection of 2  $\mu$ L). Stirring was held at 310 rpm, and injection spacing of 6 min were adopted. Calorimetric data were analyzed using MicroCal ORIGIN software (OriginLab, Northampton, MA), which uses non-linear least squares analysis to fit the data. Dilution and mixing effects were small, checked by titrating buffer into buffer, buffer into mRNA, and by observing the last injections of each experiment. All experiments were designed to include injections after saturation of the binding event to obtain baselines of the enthalpic contributions of dilution and mixing for each experiment. We corrected the

results by using the heat measured from last injections and subtracted this signal from the raw data. Origin uses the following equations for data analyses:

**Equation 4.1 Association affinity constant**

$$K = \frac{\Theta}{(1 - \Theta)[X]}$$

**Equation 4.2 Concentration of ligand X**

$$Xt = [X] + n\Theta Mt$$

$\Theta$  is the fraction of sites occupied by the ligand X

K is the association affinity constant ( $K_A$ )

$Mt$  and  $[M]$  are bulk and free concentration of macromolecule in  $V_0$ ;

$V_0$  is the active cell volume;

$Xt$  and  $[X]$  are bulk and free concentrations of ligand;

$n$  = number of sites.

Combining equations 4.1 and 4.2, gives:

**Equation 4.3 Fraction of sites occupied by the ligand X**

$$\Theta^2 - \Theta \left[ 1 + \frac{Xt}{nMt} + \frac{1}{nKMt} \right] + \frac{Xt}{nMt} = 0$$

The total heat content  $Q$  of the solution contained in  $V_0$  at fractional saturation  $\Theta$  is:

**Equation 4.4 Total heat**

$$Q = n\Theta Mt \Delta H V_0$$

$\Delta H$  is the molar heat of ligand binding. Solving the quadratic equation 4.3 for  $\Theta$  and then

substituting this into equation 4.4 gives:

**Equation 4.5 Total heat for an injection**

$$Q = \frac{nMt \Delta H V_0}{2} \left[ 1 + \frac{Xt}{nMt} + \frac{1}{nKMt} - \sqrt{\left( 1 + \frac{Xt}{nMt} + \frac{1}{nKMt} \right)^2 - \frac{4Xt}{nMt}} \right]$$

The value calculated for Q above is for the  $i^{\text{th}}$  injection in a volume  $V_0$ . Therefore, corrections are made for volume displacements with further injections. The expression for heat released  $\Delta Q(i)$ , from the  $i^{\text{th}}$  injection is:

**Equation 4.6 Heat released from a injection**

$$\Delta Q(i) = Q(i) + \frac{dVi}{V_0} \left[ \frac{Q(i) + Q(i-1)}{2} \right] - Q(i-1)$$

The process of fitting experimental data involves 1) initial guesses (which are usually accurately conducted by Origin) of  $n$ ,  $K$ , and  $\Delta H$ ; 2) calculation of  $\Delta Q(i)$  for each injection and comparison of these values with the measured heat for the corresponding experimental injection; 3) improvement in the initial values of  $n$ ,  $K$ , and  $\Delta H$  by standard Marquardt methods; and 4) iteration of the above procedure until no further significant improvement in fit occurs with continued iterations. The model for the two sets of independent sites or for the sequential binding sites are similar to the above but they add  $K_{A2}, \dots, K_{An}$  to  $K_{A1}$ .

### **Electrophoretic Mobility Shift Assays (EMSA or gel shifts)**

The procedures for EMSA have been previously described (Koslowsky et al. 2004). We 5' end-labeled the gRNA and hybridized with the cognate mRNA in the same ITC buffer containing 3% glycerol and 0.05% xylene cyanol. After denaturing and hybridizing the RNAs together for 3-4h at 27°C, samples were loaded on a 6% native polyacrylamide gel containing 2 mM  $\text{MgCl}_2$  and electrophoresed for 5 h on a running buffer also containing 2 mM  $\text{MgCl}_2$ . The gels were fixed, dried, and exposed to a phosphor screen, which was then scanned on a phosphorimager (Molecular Dynamics).

## **ACKNOWLEDGEMENTS**

This work was supported by the National Institutes of Health Grant AI34155 to D.K. We thank Dr. Charles Hoogstraten, for the VP-ITC instrument and help with the experiments. We would also like to thank Dr. Ronald Patterson, Dr. John Wang, Dr. John Fyfe, and Dr. Laura Yu for critical opinions and Dr. Karen Friderici for the NanoDrop.

## REFERENCES

- Ammerman, M.L., Fisk, J.C., and Read, L.K. 2008. gRNA/pre-mRNA annealing and RNA chaperone activities of RBP16. *RNA* **14**: 1069-1080.
- Benne, R. 1994. RNA editing in trypanosomes. *Eur J Biochem* **221**: 9-23.
- Benne, R., Van den Burg, J., Brakenhoff, J.P., Sloof, P., Van Boom, J.H., and Tromp, M.C. 1986. Major transcript of the frameshifted coxII gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA. *Cell* **46**: 819-826.
- Carnes, J., Trotter, J.R., Peltan, A., Fleck, M., and Stuart, K. 2008. RNA editing in *Trypanosoma brucei* requires three different editosomes. *Mol Cell Biol* **28**: 122-130.
- Chaires, J.B. 2001. Analysis and interpretation of ligand-DNA binding isotherms. *Methods Enzymol* **340**: 3-22.
- Cifuentes-Rojas, C., Halbig, K., Sacharidou, A., De Nova-Ocampo, M., and Cruz-Reyes, J. 2005. Minimal pre-mRNA substrates with natural and converted sites for full-round U insertion and U deletion RNA editing in trypanosomes. *Nucleic Acids Res* **33**: 6610-6620.
- Cifuentes-Rojas, C., Pavia, P., Hernandez, A., Osterwisch, D., Puerta, C., and Cruz-Reyes, J. 2006. Substrate determinants for RNA editing and editing complex interactions at a site for full-round U insertion. *J Biol Chem*.
- Cifuentes-Rojas, C., Pavia, P., Hernandez, A., Osterwisch, D., Puerta, C., and Cruz-Reyes, J. 2007. Substrate determinants for RNA editing and editing complex interactions at a site for full-round U insertion. *J Biol Chem* **282**: 4265-4276.
- Connell, G.J., Byrne, E.M., and Simpson, L. 1997. Guide RNA-independent and guide RNA-dependent uridine insertion into cytochrome b mRNA in a mitochondrial lysate from *Leishmania tarentolae*. Role of RNA secondary structure. *J Biol Chem* **272**: 4212-4218.
- Connell, G.J., and Simpson, L. 1998. Role of RNA Structure in RNA Editing. In *RNA Structure and Function*. (eds. R.W. Simons, and M. Grunberg-Manago), pp. 641-667. Cold Spring Harbor Laboratory Press, Plainview, NY.
- Cruz-Reyes, J., and Sollner-Webb, B. 1996. Trypanosome U-deletional RNA editing involves guide RNA-directed endonuclease cleavage, terminal U exonuclease, and RNA ligase activities. *Proc Natl Acad Sci USA* **93**: 8901-8906.
- Cruz-Reyes, J., Zhelonkina, A., Rusche, L., and Sollner-Webb, B. 2001. Trypanosome RNA editing: simple guide RNA features enhance U deletion 100-fold. *Mol Cell Biol* **21**: 884-892.

- Dimitrov, R.A., and Zuker, M. 2004. Prediction of hybridization and melting for double-stranded nucleic acids. *Biophys J* **87**: 215-226.
- Feig, A.L. 2007. Applications of isothermal titration calorimetry in RNA biochemistry and biophysics. *Biopolymers* **87**: 293-301.
- Golden, D.E., and Hajduk, S.L. 2006. The importance of RNA structure in RNA editing and a potential proofreading mechanism for correct guide RNA:pre-mRNA binary complex formation. *J Mol Biol* **359**: 585-596.
- Hermann, T., Schmid, B., Heumann, H., and Goring, H.U. 1997. A three-dimensional working model for a guide RNA from *Trypanosoma brucei*. *Nucleic Acids Res* **25**: 2311-2318.
- Holdgate, G.A. 2001. Making cool drugs hot: isothermal titration calorimetry as a tool to study binding energetics. *Biotechniques* **31**: 164-166, 168, 170 passim.
- Kabb, A.L., Opegard, L.M., McKenzie, B.A., and Connell, G.J. 2001. A mRNA determinant of gRNA-directed kinetoplastid editing. *Nucleic Acids Res* **29**: 2575-2580.
- Kable, M.L., Seiwert, S.D., Heidmann, S., and Stuart, K. 1996. RNA editing: a mechanism for gRNA-specified uridylate insertion into precursor mRNA. *Science* **273**: 1189-1195.
- Koslowsky, D.J., Reifur, L., Yu, L.E., and Chen, W. 2004. Evidence for U-tail stabilization of gRNA/mRNA interactions in kinetoplastid RNA editing. *RNA Biol* **1**: 28-34.
- Leung, S.S., and Koslowsky, D.J. 1999. Mapping contacts between gRNA and mRNA in trypanosome RNA editing. *Nucleic Acids Res* **27**: 778-787.
- Leung, S.S., and Koslowsky, D.J. 2001a. Interactions of mRNAs and gRNAs involved in trypanosome mitochondrial RNA editing: structure probing of an mRNA bound to its cognate gRNA. *RNA* **7**: 1803-1816.
- Leung, S.S., and Koslowsky, D.J. 2001b. RNA editing in *Trypanosoma brucei*: characterization of gRNA U-tail interactions with partially edited mRNA substrates. *Nucleic Acids Res* **29**: 703-709.
- Madison-Antenucci, S., Grams, J., and Hajduk, S.L. 2002. Editing machines: the complexities of trypanosome RNA editing. *Cell* **108**: 435-438.
- Markham, N.R., and Zuker, M. 2005. DINAMelt web server for nucleic acid melting prediction. *Nucleic Acids Res* **33**: W577-581.
- Mathews, D.H., Disney, M.D., Childs, J.L., Schroeder, S.J., Zuker, M., and Turner, D.H. 2004. Incorporating chemical modification constraints into a dynamic programming

- algorithm for prediction of RNA secondary structure. *Proc Natl Acad Sci USA* **101**: 7287-7292.
- Mikulecky, P.J., Takach, J.C., and Feig, A.L. 2004. Entropy-driven folding of an RNA helical junction: an isothermal titration calorimetric analysis of the hammerhead ribozyme. *Biochemistry* **43**: 5870-5881.
- Ochsenreiter, T., Cipriano, M., and Hajduk, S.L. 2008. Alternative mRNA editing in trypanosomes is extensive and may contribute to mitochondrial protein diversity. *PLoS ONE* **3**: e1566.
- Oppegard, L.M., Hillestad, M., McCarthy, R.T., Pai, R.D., and Connell, G.J. 2003. Cis-acting elements stimulating kinetoplastid guide RNA-directed editing. *J Biol Chem* **278**: 51167-51175.
- Pai, R.D., Oppegard, L.M., and Connell, G.J. 2003. Sequence and structural requirements for optimal guide RNA-directed insertional editing within *Leishmania tarentolae*. *RNA* **9**: 469-483.
- Piller, K.J., Decker, C.J., Rusche, L.N., Harris, M.E., Hajduk, S.L., and Sollner-Webb, B. 1995. Editing domains of *Trypanosoma brucei* mitochondrial RNAs identified by secondary structure. *Mol Cell Biol* **15**: 2916-2924.
- Russell, R. 2008. RNA misfolding and the action of chaperones. *Front Biosci* **13**: 1-20.
- Schmid, B., Riley, G.R., Stuart, K., and Goring, H.U. 1995. The secondary structure of guide RNA molecules from *Trypanosoma brucei*. *Nucleic Acids Res* **23**: 3093-3102.
- Seiwert, S.D., Heidmann, S., and Stuart, K. 1996. Direct visualization of uridylyate deletion in vitro suggests a mechanism for kinetoplastid RNA editing. *Cell* **84**: 831-841.
- Seiwert, S.D., and Stuart, K. 1994. RNA editing: transfer of genetic information from gRNA to precursor mRNA in vitro. *Science* **266**: 114-117.
- Simpson, L., Sbicego, S., and Aphasizhev, R. 2003. Uridine insertion/deletion RNA editing in trypanosome mitochondria: a complex business. *RNA* **9**: 265-276.
- Stuart, K.D., Schnauffer, A., Ernst, N.L., and Panigrahi, A.K. 2005. Complex management: RNA editing in trypanosomes. *Trends Biochem Sci* **30**: 97-105.
- Sturm, N.R., and Simpson, L. 1990. Kinetoplast DNA minicircles encode guide RNAs for editing of cytochrome oxidase subunit III mRNA. *Cell* **61**: 879-884.
- Turner, D.H. 1999. Conformational changes. In *Nucleic Acids*. (eds. V.A. Bloomfield, D.M. Crothers, and I.J. Tinoco). University Science Books, Sausalito.

- Vondruskova, E., van den Burg, J., Zikova, A., Ernst, N.L., Stuart, K., Benne, R., and Lukes, J. 2005. RNA interference analyses suggest a transcript-specific regulatory role for mitochondrial RNA-binding proteins MRP1 and MRP2 in RNA editing and other RNA processing in *Trypanosoma brucei*. *J Biol Chem* **280**: 2429-2438.
- Wiseman, T., Williston, S., Brandts, J.F., and Lin, L.N. 1989. Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal Biochem* **179**: 131-137.
- Yu, L.E., and Koslowsky, D.J. 2006. Interactions of mRNAs and gRNAs involved in trypanosome mitochondrial RNA editing: structure probing of a gRNA bound to its cognate mRNA. *RNA* **12**: 1050-1060.
- Zhuang, X., Kim, H., Pereira, M.J., Babcock, H.P., Walter, N.G., and Chu, S. 2002. Correlating structural dynamics and function in single ribozyme molecules. *Science* **296**: 1473-1476.
- Zikova, A., Horakova, E., Jirku, M., Dunajcikova, P., and Lukes, J. 2006. The effect of down-regulation of mitochondrial RNA-binding proteins MRP1 and MRP2 on respiratory complexes in procyclic *Trypanosoma brucei*. *Mol Biochem Parasitol* **149**: 65-73.
- Zikova, A., Kopečna, J., Schumacher, M.A., Stuart, K., Trantirek, L., and Lukes, J. 2008. Structure and function of the native and recombinant mitochondrial MRP1/MRP2 complex from *Trypanosoma brucei*. *Int J Parasitol* **38**: 901-912.
- Zuker, M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* **31**: 3406-3415.

## **Chapter 5**

# **CONCLUSIONS AND FUTURE DIRECTIONS**

## **INTRODUCTION**

African trypanosomes cycle between invertebrate and mammalian hosts, hence face drastic changes, such as in temperature, host immune response and energy supply. To adapt to these changes, they undergo developmental, morphological, and metabolic changes. In the mammalian host, *T. brucei* uses glucose as the main energy source and degrades it in the glycosome and cytoplasm. The mitochondrion is small and functionally repressed in this bloodstream form (BF) of the parasite (Matthews 2005). Of the 12 mRNAs that require editing, only A6, COIII, ND7, ND8, and ND9 are edited and translated in the BF, most of the mitochondrial respiratory complex proteins are down regulated (Hajduk and Sabatini 1998). In the insect host, the main energy source for the procyclic trypanosome is amino acids, which are degraded in the large and fully active mitochondrion, and the degradation products reoxidized by the mitochondrial respiratory chain (Hellemond et al. 2005). These metabolic changes are developmentally regulated and possibly due to coordinated control of nuclear- and mitochondrial-encoded proteins. Mitochondrial gene expression depends on RNA editing; however, it is unknown how editing is controlled (Koslowsky et al. 1992; Priest and Hajduk 1994). The mRNAs, gRNAs, editosome, and accessory factors are always present in the mitochondrion (Hajduk and Sabatini 1998), and annealing of the naked A6 mRNA/gRNA is sufficient to start in vitro editing in the presence of purified editing proteins (Seiwert et al. 1996; Cruz-Reyes et al. 2001). No other mRNA, except A6, will undergo a full round of editing in vitro without modifications within the sequence of the mRNA or gRNA to allow better annealing (Byrne et al. 1996; Kable et al. 1996; Adler and Hajduk 1997; Kapushoc and Simpson 1999; Kabb et al. 2001; Leung and Koslowsky 2001a; Cifuentes-

Rojas et al. 2005; Golden and Hajduk 2006). The annealing, and consequently editing efficiency, can be further improved with addition of mitochondrial accessory factors (Lambert et al. 1999; Blom et al. 2001; Muller et al. 2001; Muller and Goring 2002; Miller et al. 2006; Ammerman et al. 2008; Zikova et al. 2008). Considering the presence of hundreds of mRNA/gRNA complexes without a consensus sequence and only three editosomes (Carnes et al. 2008), we believe that RNA structure would play a role in the pathway. While a cis-element within CYb and ND7 mRNAs from *L. tarentolae* has been implicated affecting editing efficiency (Kabb et al. 2001; Oppegard et al. 2003), the mRNA secondary structure, but not sequence, was found to be a determinant of both gRNA binding affinity and editing in *T. brucei* (Koslowsky et al. 2004; Cifuentes-Rojas et al. 2005; Cifuentes-Rojas et al. 2007).

In line with the idea that RNA structure is important for the editing process, the three main chapters of this manuscript focused on RNA structure. In chapter 2, we described the structure of the A6 mRNA, gRNA and mRNA/gRNA complexes, and compared them with the CYb structures. In chapter 3, we focused on mRNA structure, by comparing structure with accessibility and annealing with the gRNA. Lastly, in chapter 4, we focused on the gRNA structure, more specifically on the effects of its U-tail on complex thermodynamics.

## **SUMMARY OF RESULTS**

### **Chapter 2 - A6 mRNA/gRNA secondary structure during editing**

In this chapter, we described the secondary structure of the complex A6 mRNA with gA6-14. Using a combination of crosslinking, solution structure probing, and computational analyses, we found it to be three-helical and conserved after the editing of

three editing sites. This conformation is similar to the structure reported for the CYb/gCYb-558 complex, supporting our main hypothesis. We additionally observed that the gA6-14 U-tail provides not only stability to the U-tail helix, but also to the other two helices within the three-helical structure. Regarding the editing site, while in the CYb complex the first editing site was of the insertion type and open to the solvent, the nucleotides at and 5' of the A6 editing sites 1 and 4 were stacked between the anchor and the U-tail helices. This local conformation within the A6 complex was surprising because we thought that the unpaired nucleotides would loop out to facilitate cleavage and editing. However, we know that between the A6 deletion sites 1 and 4 there is an insertion site, which suggests disassembly of the deletion editosome from the mRNA/gRNA complex to allow assembly of the insertion editosome. Therefore, the organization detected at ES1 and 4 could add specificity to the assembly/disassembly and cleavage steps. Distinct structural organization at the ESs could possibly distinguish an insertion from a deletion site, guiding the correct assembly with the appropriate editosome.

### **Chapter 3 - mRNA structure and gRNA targeting**

In this chapter, we showed that distinct mRNAs had their ABSs positioned within mRNAs of distinct secondary structures (distinct free energies) and the type of structure could be correlated with accessibility and binding affinity. The more structured the ABS, the more inaccessible to binding and the weaker the binding affinity. This finding was in accordance with the idea that some mRNA/gRNA complexes may need the help of annealing proteins while others may not (Vondruskova et al. 2005; Ammerman et al. 2008). The unstructured A6 and ND7-550 ABSs were very accessible and had high

binding affinity to their gRNAs. On other hand, the closed ABSs within the CYb and ND7UHR3 mRNAs correlated with lower accessibilities and lower binding affinities. The limitation imposed by accessibility and binding affinity could be involved in the developmental regulation of edited CYb and ND7UHR3 mRNAs.

#### **Chapter 4 - mRNA/gRNA binding thermodynamics**

From the results reported in chapter3, we found that distinct complexes anneal with distinct binding affinities and kinetics, which can be correlated with target structure and accessibility. Also in chapter 3, we showed that the gRNA U-tail contributed to the binding affinity of the ND7-550/gND7-550, ND7UHR3/gND7-506, CYb/gCYb-558, and A6/gA6-14 complexes. In chapter 4, we took a closer look at the A6 complex formation using 3 gRNA constructs (one with the native U-tail, one without the U-tail, and one with an enhanced U-tail). We found that with any of the constructs, gRNA binding to the A6 mRNA was an enthalpically driven process, in favor of an energetically stable bimolecular structure. The titration of gA6-14 demonstrated a greater enthalpy of binding compared with gA6-14sU. An even higher change in enthalpy was observed with gA6-14-C6, suggesting that the U-tail and the enhanced U-tail contribute to complex formation by increasing the number of hydrogen bonds.

#### ***DISCUSSION, CONCLUSIONS AND FUTURE DIRECTIONS***

Kinetoplastid RNA editing is being investigated currently by only approximately eleven laboratories in the world. Since its discovery, in 1986, much has been done to decipher the editing components, including the editosome proteins, accessory factors, the editing reactions, and potential regulatory mechanisms. Regarding the RNA components, both

the mRNAs and gRNAs involved in editing are all of distinct primary structures and share no common sequence, except for the gRNA 3' U-tail (Blum et al. 1990; Piller et al. 1995; Leung and Koslowsky 1999). Three editosomes have been characterized, one for insertion, one for deletion and another for editing in cis (Carnes et al. 2008). Their specificity and efficiency are dependent on the structure and stability of the mRNA/gRNA complex (Cifuentes-Rojas et al. 2007; Hernandez et al. 2008), but not much is known about the distinct mRNA/gRNA complexes.

We hypothesize that mRNA/gRNA annealing may be a limiting factor in complex formation. In addition, once the RNAs are annealed all the different mRNA/gRNA complexes must present common structural features that are important for editosome recognition. Previous predictions coupled with solution structure probing of the CYb mRNA/gRNA complex indicate that it folds into three helices that are maintained as editing proceeds (Blum and Simpson 1990; Leung and Koslowsky 2001a; Yu and Koslowsky 2006). In chapter 2, we described that the structure for the A6 pair is also three-helical and persists through editing of the initial editing sites. This finding greatly supports our main hypothesis of a core structure coexisting among distinct mRNA/gRNA complexes. The biological importance of such conformation remains to be determined. Three- and four-helical junctions are common in a broad range of functional RNAs. Four-way junctions can be found in rRNA, U1 snRNA, and in the hairpin ribozyme of the satellite RNA of the tobacco ringspot virus. Three-way junctions are present also in rRNA, in P4-P6 domain of the group I intron ribozyme and in the hammerhead ribozyme of plant viroids. In addition to serving as protein binding sites, the junctions have been shown to provide flexibility for tertiary interaction, allowing for the existence of

alternative folding (Lilley 1998; 2000). When the alternative conformer is stable it can control the RNA function or the RNA interaction with proteins, and hence regulate the system in which it is involved (Lilley 2000; Glazov et al. 2006). Thus, the purpose for the the various mRNA/gRNA complexes to fold into a common structure could be simply to be recognized by the editosome or, additionally, the complex could further fold into an alternative conformer to regulate editing. With the work presented in this manuscript, we confirmed that distinct mRNA/gRNA complexes can fold into a common core structure. In addition, we found that the core structure presents pre-formed structural differences at the editing site that could allow for distinction between insertion and deletion sites.

gRNAs are always present in the mitochondria of BF and PF, even when a specific gRNA-binding protein is knocked down. Thus, the gRNA abundance is not implicated in editing control (Koslowsky et al. 1992; Lambert et al. 1999). Therefore, it has been suggested that editing regulation may be through distinct binding affinities of gRNAs to their cognate mRNAs. In addition to checking the binding affinity of distinct mRNA/gRNA complexes, in chapter 3 of this manuscript we determined the accessibility of distinct mRNAs to the respective gRNAs. Target structure has been extensively demonstrated to have a powerful effect on accessibility and efficiency of several RNA systems, including the binding pathways for miRNA, siRNA, ribozymes, and antisense RNA or ODN (Lima et al. 1992; Ho et al. 1998; Amarzguioui et al. 2000; Vickers et al. 2000; Long et al. 2007; Shao et al. 2007a; Shao et al. 2007b). Here, we found that the anchor binding sites within distinct mRNAs had several levels of accessibility that were dependent on the surrounding structure. mRNAs with high accessibility directly correlated to a stronger binding affinity for the cognate gRNA and to some extent, to

editing efficiency. In addition, the gRNA U-tail seemed to improve the binding affinity of some but not all mRNA/gRNA complexes. Though editing control is likely to be more complex and involve mitochondrial proteins, we demonstrated that mRNA structure can disturb mRNA/gRNA complex formation by itself, serving as a limiting step on the editing pathway. To complement these experiments with purified RNAs, it will be interesting to conduct the same accessibility and binding affinity experiments in the presence of the mitochondrial extract.

The gRNA U-tail is the only common motif within the RNAs involved in editing. It has been shown to improve mRNA/gRNA stability, binding affinity, and editing efficiency to different degrees in *T. brucei* (Blum and Simpson 1990; Burgess et al. 1999; Kapushoc and Simpson 1999; Cruz-Reyes et al. 2001; Koslowsky et al. 2004). To understand the U-tail function in complex formation, our goal was to determine the thermodynamic parameters of this bimolecular interaction. The systematic nearest neighbor studies of bimolecular interaction are unable to account for tertiary contacts and flanking helices (Dimitrov and Zuker 2004) and optical melting is only accurate for systems that undergo two-state folding. Since the mRNA/gRNA complex involves tertiary contacts, forms a flanking helix (the gRNA stem-loop), and is unlikely to undergo two-state folding (Schmid et al. 1996) we opted to use ITC. Complex formation was energetically favored, independently of the presence/absence of the U-tail, or in the presence of an enhanced U-tail, as expected. The U-tail and the enhanced U-tail are thought to improve the complex stability and binding affinity, as predicted by computer algorithms or measured by EMSA (Cruz-Reyes et al. 2001; Koslowsky et al. 2004). Unfortunately, the presence of the tail in gA6-14 caused a distortion on the ITC data,

resulting in poor curve fitting that is difficult to interpret. We believe that alternative (kinetically trapped) conformers probably caused the poor ITC curve fit for the gA6-14 and gA6-14-C6 experiments. Deleting the initial titrations for gA6-14 and gA6-14-C6 (caused by unspecific reactions) showed that the data is comparable to gA6-14sU. From these, we can tell that the presence of the U-tail within gA6-14 promotes a higher change in enthalpy, caused by its interaction with the mRNA. Since ITC directly measures binding thermodynamics in solution and overlaps the energy from structure disruption, nucleation, hybridization, and the tertiary components of folding (Feig 2007) the values obtained by ITC reflect the energetics of binding under biologically relevant conditions. Conducting ITC at higher temperatures and salt conditions may contribute to the understanding of structure playing a role in complex thermodynamics for the gRNAs with a tail.

Taken together, the work presented here contributes to the understanding of steps prior and subsequent to mRNA/gRNA complex formation. We have confirmed that the structure of the mRNA and gRNA interferes with the binding event, and that distinct mRNA complexes fold into an energetically favorable three-helical structure. With the discovery of annealing proteins, future experimental projects should take these into consideration and the editosome proteins, with the purpose of evaluating how structure and binding thermodynamics are altered in their presence.

## REFERENCES

- Adler, B.K., and Hajduk, S.L. 1997. Guide RNA requirement for editing-site-specific endonucleolytic cleavage of preedited mRNA by mitochondrial ribonucleoprotein particles in *Trypanosoma brucei*. *Mol Cell Biol* **17**: 5377-5385.
- Amarzguioui, M., Brede, G., Babaie, E., Grotli, M., Sproat, B., and Prydz, H. 2000. Secondary structure prediction and in vitro accessibility of mRNA as tools in the selection of target sites for ribozymes. *Nucleic Acids Res* **28**: 4113-4124.
- Ammerman, M.L., Fisk, J.C., and Read, L.K. 2008. gRNA/pre-mRNA annealing and RNA chaperone activities of RBP16. *RNA* **14**: 1069-1080.
- Blom, D., Burg, J., Breek, C.K., Speijer, D., Muijsers, A.O., and Benne, R. 2001. Cloning and characterization of two guide RNA-binding proteins from mitochondria of *Crithidia fasciculata*: gBP27, a novel protein, and gBP29, the orthologue of *Trypanosoma brucei* gBP21. *Nucleic Acids Res* **29**: 2950-2962.
- Blum, B., Bakalara, N., and Simpson, L. 1990. A model for RNA editing in kinetoplastid mitochondria: "guide" RNA molecules transcribed from maxicircle DNA provide the edited information. *Cell* **60**: 189-198.
- Blum, B., and Simpson, L. 1990. Guide RNAs in kinetoplastid mitochondria have a nonencoded 3' oligo(U) tail involved in recognition of the preedited region. *Cell* **62**: 391-397.
- Burgess, M.L., Heidmann, S., and Stuart, K. 1999. Kinetoplastid RNA editing does not require the terminal 3' hydroxyl of guide RNA, but modifications to the guide RNA terminus can inhibit in vitro U insertion. *RNA* **5**: 883-892.
- Byrne, E.M., Connell, G.J., and Simpson, L. 1996. Guide RNA-directed uridine insertion RNA editing in vitro. *Embo J* **15**: 6758-6765.
- Carnes, J., Trotter, J.R., Peltan, A., Fleck, M., and Stuart, K. 2008. RNA editing in *Trypanosoma brucei* requires three different editosomes. *Mol Cell Biol* **28**: 122-130.
- Cifuentes-Rojas, C., Halbig, K., Sacharidou, A., De Nova-Ocampo, M., and Cruz-Reyes, J. 2005. Minimal pre-mRNA substrates with natural and converted sites for full-round U insertion and U deletion RNA editing in trypanosomes. *Nucleic Acids Res* **33**: 6610-6620.
- Cifuentes-Rojas, C., Pavia, P., Hernandez, A., Osterwisch, D., Puerta, C., and Cruz-Reyes, J. 2007. Substrate determinants for RNA editing and editing complex interactions at a site for full-round U insertion. *J Biol Chem* **282**: 4265-4276.

- Cruz-Reyes, J., Zhelonkina, A., Rusche, L., and Sollner-Webb, B. 2001. Trypanosome RNA editing: simple guide RNA features enhance U deletion 100-fold. *Mol Cell Biol* **21**: 884-892.
- Dimitrov, R.A., and Zuker, M. 2004. Prediction of hybridization and melting for double-stranded nucleic acids. *Biophys J* **87**: 215-226.
- Feig, A.L. 2007. Applications of isothermal titration calorimetry in RNA biochemistry and biophysics. *Biopolymers* **87**: 293-301.
- Glazov, E.A., Pheasant, M., Nahkuri, S., and Mattick, J.S. 2006. Evidence for control of splicing by alternative RNA secondary structures in Dipteran homothorax pre-mRNA. *RNA Biol* **3**: 36-39.
- Golden, D.E., and Hajduk, S.L. 2006. The importance of RNA structure in RNA editing and a potential proofreading mechanism for correct guide RNA:pre-mRNA binary complex formation. *J Mol Biol* **359**: 585-596.
- Hajduk, S., and Sabatini, R.S. 1998. Mitochondrial mRNA editing in kinetoplastid protozoa. In *Modification and editing of RNA*. (eds. H. Grosjean, and R. Benne). ASM Press, Washington, DC.
- Hellemond, J.J., Bakker, B.M., and Tielens, A.G. 2005. Energy metabolism and its compartmentation in *Trypanosoma brucei*. *Adv Microb Physiol* **50**: 199-226.
- Hernandez, A., Panigrahi, A., Cifuentes-Rojas, C., Sacharidou, A., Stuart, K., and Cruz-Reyes, J. 2008. Determinants for association and guide RNA-directed endonuclease cleavage by purified RNA editing complexes from *Trypanosoma brucei*. *J Mol Biol* **381**: 35-48.
- Ho, S.P., Bao, Y., Leshner, T., Malhotra, R., Ma, L.Y., Fluharty, S.J., and Sakai, R.R. 1998. Mapping of RNA accessible sites for antisense experiments with oligonucleotide libraries. *Nat Biotechnol* **16**: 59-63.
- Kabb, A.L., Opegard, L.M., McKenzie, B.A., and Connell, G.J. 2001. A mRNA determinant of gRNA-directed kinetoplastid editing. *Nucleic Acids Res* **29**: 2575-2580.
- Kable, M.L., Seiwert, S.D., Heidmann, S., and Stuart, K. 1996. RNA editing: a mechanism for gRNA-specified uridylyate insertion into precursor mRNA. *Science* **273**: 1189-1195.
- Kapushoc, S.T., and Simpson, L. 1999. In vitro uridine insertion RNA editing mediated by cis-acting guide RNAs. *RNA* **5**: 656-669.
- Koslowsky, D.J., Reifur, L., Yu, L.E., and Chen, W. 2004. Evidence for U-tail stabilization of gRNA/mRNA interactions in kinetoplastid RNA editing. *RNA Biol* **1**: 28-34.

- Koslowsky, D.J., Riley, G.R., Feagin, J.E., and Stuart, K. 1992. Guide RNAs for transcripts with developmentally regulated RNA editing are present in both life cycle stages of *Trypanosoma brucei*. *Mol Cell Biol* **12**: 2043-2049.
- Lambert, L., Muller, U.F., Souza, A.E., and Goring, H.U. 1999. The involvement of gRNA-binding protein gBP21 in RNA editing-an in vitro and in vivo analysis. *Nucleic Acids Res* **27**: 1429-1436.
- Leung, S.S., and Koslowsky, D.J. 1999. Mapping contacts between gRNA and mRNA in trypanosome RNA editing. *Nucleic Acids Res* **27**: 778-787.
- Leung, S.S., and Koslowsky, D.J. 2001a. Interactions of mRNAs and gRNAs involved in trypanosome mitochondrial RNA editing: structure probing of an mRNA bound to its cognate gRNA. *RNA* **7**: 1803-1816.
- Lilley, D.M. 1998. Folding of branched RNA species. *Biopolymers* **48**: 101-112.
- Lilley, D.M. 2000. Structures of helical junctions in nucleic acids. *Q Rev Biophys* **33**: 109-159.
- Lima, W.F., Monia, B.P., Ecker, D.J., and Freier, S.M. 1992. Implication of RNA structure on antisense oligonucleotide hybridization kinetics. *Biochemistry* **31**: 12055-12061.
- Long, D., Lee, R., Williams, P., Chan, C.Y., Ambros, V., and Ding, Y. 2007. Potent effect of target structure on microRNA function. *Nat Struct Mol Biol* **14**: 287-294.
- Matthews, K.R. 2005. The developmental cell biology of *Trypanosoma brucei*. *J Cell Sci* **118**: 283-290.
- Miller, M.M., Halbig, K., Cruz-Reyes, J., and Read, L.K. 2006. RBP16 stimulates trypanosome RNA editing in vitro at an early step in the editing reaction. *RNA* **12**: 1292-1303.
- Muller, U.F., and Goring, H.U. 2002. Mechanism of the gBP21-mediated RNA/RNA annealing reaction: matchmaking and charge reduction. *Nucleic Acids Res* **30**: 447-455.
- Muller, U.F., Lambert, L., and Goring, H.U. 2001. Annealing of RNA editing substrates facilitated by guide RNA-binding protein gBP21. *Embo J* **20**: 1394-1404.
- Oppegard, L.M., Hillestad, M., McCarthy, R.T., Pai, R.D., and Connell, G.J. 2003. Cis-acting elements stimulating kinetoplastid guide RNA-directed editing. *J Biol Chem* **278**: 51167-51175.
- Piller, K.J., Decker, C.J., Rusche, L.N., Harris, M.E., Hajduk, S.L., and Sollner-Webb, B. 1995. Editing domains of *Trypanosoma brucei* mitochondrial RNAs identified by secondary structure. *Mol Cell Biol* **15**: 2916-2924.

- Priest, J.W., and Hajduk, S.L. 1994. Developmental regulation of mitochondrial biogenesis in *Trypanosoma brucei*. *J Bioenerg Biomembr* **26**: 179-191.
- Schmid, B., Read, L.K., Stuart, K., and Goring, H.U. 1996. Experimental verification of the secondary structures of guide RNA-pre-mRNA chimaeric molecules in *Trypanosoma brucei*. *Eur J Biochem* **240**: 721-731.
- Seiwert, S.D., Heidmann, S., and Stuart, K. 1996. Direct visualization of uridylate deletion in vitro suggests a mechanism for kinetoplastid RNA editing. *Cell* **84**: 831-841.
- Shao, Y., Chan, C.Y., Maliyekkel, A., Lawrence, C.E., Roninson, I.B., and Ding, Y. 2007a. Effect of target secondary structure on RNAi efficiency. *RNA* **13**: 1631-1640.
- Shao, Y., Wu, S., Chan, C.Y., Klapper, J.R., Schneider, E., and Ding, Y. 2007b. A structural analysis of in vitro catalytic activities of hammerhead ribozymes. *BMC Bioinformatics* **8**: 469.
- Vickers, T.A., Wyatt, J.R., and Freier, S.M. 2000. Effects of RNA secondary structure on cellular antisense activity. *Nucleic Acids Res* **28**: 1340-1347.
- Vondruskova, E., van den Burg, J., Zikova, A., Ernst, N.L., Stuart, K., Benne, R., and Lukes, J. 2005. RNA interference analyses suggest a transcript-specific regulatory role for mitochondrial RNA-binding proteins MRP1 and MRP2 in RNA editing and other RNA processing in *Trypanosoma brucei*. *J Biol Chem* **280**: 2429-2438.
- Yu, L.E., and Koslowsky, D.J. 2006. Interactions of mRNAs and gRNAs involved in trypanosome mitochondrial RNA editing: structure probing of a gRNA bound to its cognate mRNA. *RNA* **12**: 1050-1060.
- Zikova, A., Kopecna, J., Schumacher, M.A., Stuart, K., Trantirek, L., and Lukes, J. 2008. Structure and function of the native and recombinant mitochondrial MRP1/MRP2 complex from *Trypanosoma brucei*. *Int J Parasitol* **38**: 901-912.

MICHIGAN STATE UNIVERSITY LIBRARIES



3 1293 03062 4260