

Minimal pre-mRNA substrates with natural and converted sites for full-round U insertion and U deletion RNA editing in trypanosomes

Catherine Cifuentes-Rojas, Kari Halbig, Anastasia Sacharidou,
Monica De Nova-Ocampo and Jorge Cruz-Reyes*

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

Received September 9, 2005; Revised and Accepted October 17, 2005

ABSTRACT

Trypanosome RNA editing by uridylyate insertion or deletion cycles is a mitochondrial mRNA maturation process catalyzed by multisubunit complexes. A full-round of editing entails three consecutive steps directed by partially complementary guide RNAs: pre-mRNA cleavage, U addition or removal, and ligation. The structural and functional composition of editing complexes is intensively studied, but their molecular interactions in and around editing sites are not completely understood. In this study, we performed a systematic analysis of distal RNA requirements for full-round insertion and deletion by purified editosomes. We define minimal substrates for efficient editing of A6 and CYb model transcripts, and established a new substrate, RPS12. Important differences were observed in the composition of substrates for insertion and deletion. Furthermore, we also showed for the first time that natural sites can be artificially converted in both directions: from deletion to insertion or from insertion to deletion. Our site conversions enabled a direct comparison of the two editing kinds at common sites during substrate minimization and demonstrate that all basic determinants directing the editosome to carry out full-round insertion or deletion reside within each editing site. Surprisingly, we were able to engineer a deletion site into CYb, which exclusively undergoes insertion in nature.

INTRODUCTION

Maturation of mitochondrial pre-mRNAs in kinetoplastid species including *Trypanosoma* and the evolutionarily distant *Leishmania* involves three-step cycles of uridylyate insertion or deletion at many editing sites (ESs). Multisubunit editing complexes switch between insertion and deletion modes as editing progresses with a general 3'-to-5' polarity. The multisubunit composition and function of editing complexes is currently under intense study [for reviews see (1–3)], and their interactions with editing substrates have not been identified. Each editing cycle, directed by partially complementary mitochondrial guide RNAs (gRNAs), entails three basic consecutive catalytic steps: first, pre-mRNA cleavage; second, 3' terminal processing of the upstream piece by either nucleotide addition mediated by terminal U transferase 'TUTase' (in insertion), or nucleotide removal by a U-specific exonuclease (in deletion); finally, mRNA resealing by ligation (4–7). Apart from these basic catalytic steps, annealing and unwinding steps are also likely involved (8,9). The most evident landmarks of deletion or insertion ESs are unpaired mRNA uridylyates or unpaired gRNA purines, respectively. Site-specific mutagenesis affecting the ES or adjoining residues, particularly their potential for mRNA/gRNA pairing can significantly impact the specificity and efficiency of full-round and partial ('pre-cleaved') editing (10–12). In trypanosomes, it is generally accepted that natural ESs lie between two flanking duplexes: a proximal upstream duplex formed by a pre-mRNA 5' purine-rich and gRNA 3' poly(U) sequences (4,7,13,14), and an adjacent pre-mRNA/gRNA downstream 'anchor' duplex that directs cleavage (4–7). Site-specific mutagenesis of the gRNA 3' region, to artificially stabilize the upstream duplex, can enhance full-round editing *in vitro*

*To whom correspondence should be addressed. Tel: +1 979 458 3375; Fax: +1 979 862 4718; Email: cruzrey@tamu.edu

Present address:

Nova-Ocampo, Instituto de Ciencias Agropecuarias, Universidad Autonoma del Estado de Hidalgo. Av. Universidad Km 1, Ex-Hacienda de Aquetzalpa, Rancho Universitario. Tulancingo, Hidalgo, CP 43600, Mexico

© The Author 2005. Published by Oxford University Press. All rights reserved.

The online version of this article has been published under an open access model. Users are entitled to use, reproduce, disseminate, or display the open access version of this article for non-commercial purposes provided that: the original authorship is properly and fully attributed; the Journal and Oxford University Press are attributed as the original place of publication with the correct citation details given; if an article is subsequently reproduced or disseminated not in its entirety but only in part or as a derivative work this must be clearly indicated. For commercial re-use, please contact journals.permissions@oxfordjournals.org

(10,15). In *Leishmania*, the need for such upstream duplex remains somewhat controversial (16,17).

Trypanosome full-round deletion and insertion are currently studied *in vitro* using substrates representing fragments of the pre-edited domain in two mRNAs encoding ATPase subunit 6 (A6) and apo-cytochrome *b* (CYb) (7,11,15,18). The natural A6 pre-mRNA contains numerous insertion and deletion ESs, but only ES1 and ES2 are currently used to analyze full-round deletion and insertion *in vitro*, respectively. A6 substrates between 65 and 72 nt long have been used (7,11,18). In contrast, CYb pre-mRNA exclusively contains insertion ESs, of which only ES1 has been analyzed *in vitro* using a CYb 55 nt substrate (15). The influence of pre-mRNA structural properties, such as substrate composition, folding and potential *cis*-elements on full-round insertion and deletion, has not been systematically characterized in trypanosomes. This is of interest because the two kinds of editing in a common complex may be differentially regulated, as they involve separate activities and enzymes, and there is evidence for their physical separation in the complex (19–23). Also, different gRNA features appear to be required for deletion and insertion (10,12,24). Thus, it is feasible that editing complex interactions with a processing site and its environs differs between the two editing types.

Here, we have analyzed the effect of substrate minimization on full-round insertion and deletion at natural and converted sites in different substrates. Our studies comparing catalysis by moderately and significantly enriched editing complexes show that, at least for some substrates, the latter are more sensitive to substrate minimization. This suggests that one or more associated RNA-binding activities are sensitive to the overall size of mRNAs, and also to stringent purification of the complexes. Substrate minimization had significantly distinct effects on insertion and deletion, and the extent and direction of such effects varied among substrates. This implies that editing complexes are quite sensitive to RNA context. We also show for first time that natural sites can be converted in either direction, from deletion to insertion and *vice versa*. Such site conversion not only allowed comparison of both types of editing at specific sites during substrate minimization, but also underscores the basic nature of the RNA determinants that direct complexes to carry out full-round insertion or deletion.

MATERIALS AND METHODS

Pre-mRNA and gRNA substrates

The ATPase 6 (A6) and CYb (CYb) pre-mRNA substrates used for full-round insertion at ES2 and ES1, respectively, are derived from the A6AC and CYb anchor substrates, and enhanced gRNAs, originally described by Igo *et al.* (11,15). gCYb[558]USD was modified to direct +3U insertion, and all A6 mRNA constructs used pre-edited ES1 to increase the stability of the downstream duplex. All RNAs were transcribed using the Uhlenbeck single-stranded T7 transcription method (25) and gel purified. Converted ESs in A6, CYb and RPS12 were engineered on natural pre-mRNA sequences paired with enhanced gRNAs (10,11,15). All DNA templates below (except for PCR primers) are 3' extended with the T7 promoter complementary strand: TATAGTGAGTCGTATTA. The number of each RNA product is indicated.

(1) GATGCCAGGTAAGTATTCTATAACTCCAAAAT - CAGTACTTTCCCTTTCTTCTCTCCTCCCCCTAACCTT-TCC; (2) GGAAAGGUUAGATGCCAGGTAAGTATTCT-ATAACTCCAAAATCAGTACTTTCCCTTTCTTCTCTCC-TCCCC; (3) CTATAACTCCAAAATCAGTACTTTCCC-TTTCTTCTCTCCTCCCCCTAACCTTTCC; (4) CTATAA-CTCCAAAATCAGTACTTTCCCTTTCTTCTCTCCTCCC-CC; (5) CTATAACTCCAAAATCAGTACTTTCCCTTTC-TTCTCTCCTCCC; (6) CTATAACTCCAAAATCAGTACT-TTCCCTTTCTTCTCTCCC; (7) CTATAACTCCAAAAT-CAGTACTTTCCCTTTCTTCTCC; (8) CTATAACTCCAA-AATCAGTACTTTCCCTTTCTTC; (9) CTATAACTCCA-AAATCAGTACTTTCCCTTTCTTCAAAGAAAGAGCCC; (10) GATGCCAGGTAAGTATTCTATAACTCCAAAATC-AGTACTTTCCCTTTCTTC; (11) GATGCCAGGTAAGTA-TTCTATAACTCCAAAATAAACAGTACTTTCCCTTTCT-TCTCTCCTCCCCCTAACCTTTCC; (12) GATGCCAGG-TAAGTATTCTATAACTCCAAAATAAACAGTACTTTCC-CTTTTCTTCTCTCCTCCCC; (13) CTATAACTCCAAA-ATAAACAGTACTTTCCCTTTCTTCTCTCCTCCCCCTA-ACCTTTCC; (14) CTATAACTCCAAAATAAACAGTAC-TTTCCCTTTCTTCTCTCCTCCCC; (15) CTATAACTC-CAAATAAACAGTACTTTCCCTTTCTTCTCTCCTCCCC-CCACAAAACAAACGCGC; (16) ACACAACAACA-ACACACCTATAACTCCAAAATAAACAGTACTTTCCCT-TTTCTTCTCTCCTCCCC; (17) For.primers: ATTTAGG-TGACACTATAGTTAAGAATAATGGTTATAAATTTTA. Rev.primers: CTGACATTAAGACCCCTTTC; (18) For.primers: TAATACGACTCACTATAGAATAAATTTTATA-TAAAAGCG, Rev.primers: CTGACATTAAGACCCCT-TTC; (19) For.primers: TAATACGACTCACTATAGAAA-ATATAAAGCGG, Reverse primer: same as above; (20) CTGACATTAAGACCCCTTTCTTTTTTCTCCGCTTTT-ATATTC; (21) CTGACATTAAGACCCCTTTCTTTTT-TCTCCGCTTTTATC; (22) CTGACATTAAGACCCCTT-TTTTTTCTCCGCTTTC; (23) CTGACATTAAGAA-CAAACCTTTCTTTTTTCTCCGCTTTTATATAAAATT-TATAACCATTATTCTTTCC; (24) For.primers: TAATACG-ACTCACTATAGAATAAATTTTATATAAAGCGG, Rev.primers: CTGACATTAAGACAAAACCTTTCTTTTTTTC; (25) For.primers: TAATACGACTCACTATAGAAAATA-TAAAAGCGG, Reverse primer: same as above; (26) CTGA-CATTAAGACAAAACCTTTCTTTTTTCTCCGCTTTT-CTC; (27) CTGACATTAAGACAAAACCTTTCTTTTT-TTCTCCGCTTC; (28) CTGACATTAAGACAAAACC-TTTCTTTTTTCTCCC; (29) TAAAACATATCTTATCCT-AAAATCTTCTTAAACGATGTTTCTTTAAACCAAAC-AAAGGGTTCCCCACC; (30) Forward primer: GTA-ATACGACTCACTATAGGAATAAAGG, Reverse primer: TAAAACTAATCTTATCCTAAAATCTCTTC. DNA template: sheared DNA clone 7H20 was a kind gift from Najib M.EL.Sayed. (31) TAAAACATATCTTATCCTTCG-CTTCTAAACGATGTTTCTTCTCCTATATGGAGTCGT-ATTA.D43(0,4)ES2:AAAGAAAGGGAAAACAAAAGAG-TCTCGGAGTTATAGTATATCC; gCYb.Ins3:AAAGCGG-AGAAAAAATTCACATTTGTCTTTTAAATGTCAGTCCCT-ATAGTGAGTCGTATTA; gCYb.D39.1GGAGAAAAAGAA-AAAAGTGTCCTTTTAAATGTCAGTCCCTATAGTGAGT-CGTATTA; gRPS12.D39:ACATCGTTTTAGAAAAAGGG-TAGGATAAGATATATTTACC; gRPS12 I42.4:GAAAGA-TCGTTTTAAAATTAATTTAGGATAAGATATATTTACC.

Preparation of crude mitochondrial extracts and enrichment of editing complexes

Procyclic *Trypanosoma brucei* strain TREU667 was grown in Cunningham media, and mitochondrial extracts were prepared as described previously (26,27). Mitochondrial crude extract purified by Q-Sepharose ion exchange chromatography, or further purified by DNA-cellulose ion exchange chromatography to seven-major silver-stained components of editing complexes, was described previously (28). The elution fractions with the peak of editing activity were used in each case as indicated in the text (Q fractions and D fractions). The differences in activity by Q and D fractions described in this study were reproduced.

Editing and quantification analysis

Full-round U deletion and U insertion were performed as described by Cruz-Reyes *et al.* (19). Briefly, 3' end-labeled pre-mRNA (~10 fmol) and gRNA (~1.25 pmol) were pre-annealed and incubated in 10 mM-MRB buffer [10 mM Mg(OAc)₂, 10 mM KCl, 1 mM EDTA, pH 8, 25 mM Tris-HCl, pH 8 and 5% glycerol]. Insertion reactions were supplemented with 150 μM UTP and 3 μM ATP. Deletion reactions were supplemented with 3 mM AMP-CP and 30 μM ATP. Aliquots containing 20 μl editing reactions, including 2 μl of peak editing fraction, were incubated at 26°C for 60 min. Upon incubation, the RNA was deproteinized and resolved in 9% acrylamide, 8 M urea gels. Data were visualized by phosphorimaging and/or X-ray autoradiography. Quantification of editing products was performed using a STORM PhosphorImager (Image Quant 5.0, Molecular Dynamics). Each panel in these figures corresponds to one of two replica series performed simultaneously (i.e. one experiment). At least two independent experiments were performed for each figure and the data shown is representative. The editing activity varied between protein preparations, but the relative efficiency of the constructs was consistently observed. The abundance of accurately edited product for each construct tested was initially calculated as the percentage of total input RNA and then normalized to the abundance of the corresponding product with the most active substrate identified. Mean and error bars were plotted on a linear scale.

RESULTS

A6 insertion and deletion editing

The influence that pre-mRNAs features may have on editing efficiency, including potential *cis*-elements, substrate composition and RNA folding, have not been systematically studied. We began analyzing the contribution of upstream and downstream pre-mRNA features, by performing gradual terminal truncations on an A6 72 nt substrate (RNA1, which includes the 3'-most 26 ESs) paired to a gRNA that directs 3U insertion at the natural ES2 (Figure 1A). In these studies, we used catalytic complexes that were enriched by Q-Sepharose ion-exchange chromatography ('Q fraction'), or complexes that were significantly purified by consecutive Q-Sepharose and DNA-cellulose chromatography ('D fraction'), as previously reported (28).

Deletions of 10 nt from the 5' end (62 nt RNA2), 17 nt from the 3' end (55 nt RNA3) or both combined (45 nt RNA4)

supported accurate 3U insertion, with both Q and D fractions (Figure 1B, lanes 1–4 and 5–8, respectively). Quantification of accurate 3U insertion products showed that editing with the 45 nt RNA4 was ~2-fold more efficient than with the parental 72 nt RNA1 (Figure 1C and E). The abundance of accurately edited product for each construct tested, was initially calculated as the percentage of total input RNA and then normalized to the abundance of the corresponding product with the most active substrate identified. Mean and error bars from two experiments were plotted on a linear scale.

The 45 nt RNA4 has 30 nt (including 27 purines) upstream and 15 nt downstream of ES2. Interestingly, progressive 5' truncations of this substrate (43 nt RNA5, 40 nt RNA6, 37 nt RNA7 and 34 nt RNA8) showed quantitatively different inhibitory effects with the Q and D fractions. 3U insertion was moderately reduced with decreased template length in the Q fraction (Figure 1D, lanes 1–5; and Figure 1C). However, with the D fraction, increased truncation of the template had a more dramatic effect (Figure 1D, lanes 6–10; and Figure 1E), and no 3U insertion was detected with the shortest 34 nt template (Figure 1D, lane 10; and Figure 1E). Notably, the 45 nt RNA4 and shorter constructs are capable of forming the same upstream and downstream duplexes with gRNA and exhibit the same ΔG (Figure 1A), and the observed gradual drop in insertion parallels the extent of upstream sequence truncation. Thus, an ~45 nt A6 substrate is optimal for insertion at ES2 in both Q and D fractions, and the significant inhibition of insertion with shorter constructs may reflect, at least in part, loss of stimulatory *cis*-features (e.g. purine-richness) or a more general property, such as overall length or folding.

To distinguish between these possibilities we extended the 34 nt RNA8, that is inactive in the D fraction, either with an artificial upstream pyrimidine run (47 nt RNA9) or its natural downstream pyrimidine-rich sequence (51 nt RNA10). Note that the size of these two constructs is similar to the efficient 45 nt RNA4, although their predicted structure is 3- to 4-fold more stable (see ΔG s in Figure 1A). Both pyrimidine-rich extensions reestablished editing to at least 50% of the maximal level with RNA4 (Figure 1F and G). This suggests that reduced purine-richness was not the cause of substrate inactivation. Consistent with this notion, a 3' polypurine extension reestablished editing to a level comparable with the pyrimidine-extensions (data not shown).

Interestingly, shortening of the A6 substrate caused significant reduction in the ratio of 3U to 1U insertion (Figure 1D). In the case of the smallest 34 nt RNA8, 1U insertion was the major product in the Q fraction (lane 5), and the only detectable product in the D fraction (lane 10). This effect is more evident in the Q fraction. We are currently addressing the mechanistic reasons for this apparent reduction in insertion fidelity.

Combined, the data above indicate that for *in vitro* A6 insertion at ES2, an ~45 nt long substrate is optimal, and neither specific composition (upstream and downstream of the 34 nt RNA8 sequence) nor potential pre-mRNA folding prior to gRNA pairing appears critical. The particularly dramatic inhibitory effect of template shortening in the D fraction may be due to loss of relevant protein factors in the purified editing complexes.

A

Pre- mRNA No. nt. ΔG	A6 INSERTION	$\Delta G = -20.5$	ES2	$\Delta G = -21.5$
1 72 -13.1	GGAAAGGUUAGGGGGAGGAGAGAAGAAAGGGAAAGUACUGAUUUUUGGAGUUUAGAAUACUUACCUGGCAUC	UUUCUUUCCUUUG	UAGAGCCUCAUAUCAUAUAGG	AAGCAA
2 62 -9.7	GGGGGAGGAGAGAAGAAAGGGAAAGUACUGAUUUUUGGAGUUUAGAAUACUUACCUGGCAUC			
3 55 -4.8	GGAAAGGUUAGGGGGAGGAGAGAAGAAAGGGAAAGUACUGAUUUUUGGAGUUUAG			
4 45 -3.2	GGGGGAGGAGAGAAGAAAGGGAAAGUACUGAUUUUUGGAGUUUAG			
5 43 -3.2	GGGAGGAGAGAAGAAAGGGAAAGUACUGAUUUUUGGAGUUUAG			
6 40 -3.2	GGGAGAGAAGAAAGGGAAAGUACUGAUUUUUGGAGUUUAG			
7 37 -3.2	GGAGAAGAAAGGGAAAGUACUGAUUUUUGGAGUUUAG			
8 34 -3.2	GAAGAAAGGGAAAGUACUGAUUUUUGGAGUUUAG			
9 47 -12.9	GGCUCUUUUCUUUGAAGAAAGGGAAAGUACUGAUUUUUGGAGUUUAG			
10 51 -9.9	GAAGAAAGGGAAAGUACUGAUUUUUGGAGUUUAGAAUACUUACCUGGCAUC			

$\Delta G = -18.0$
3'duplex

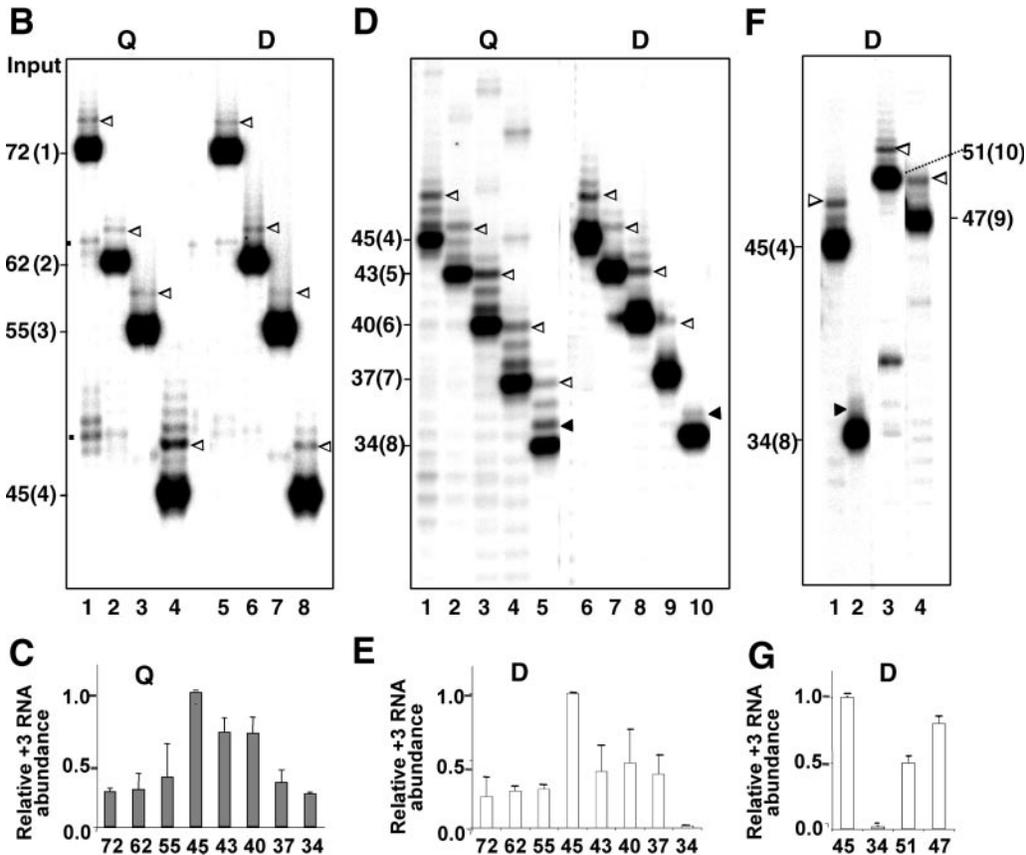


Figure 1. Mutagenesis of A6 pre-mRNA for full-round insertion at ES2. (A) The starting 72 nt RNA1 is a derivative of the reported 53 nt A6AC substrate (15) bearing a 3' natural extension and pre-edited ES1. The derived constructs used in this study are aligned underneath. Purines are in black and pyrimidines in red. The gRNA gA6[14]USD-3A [(15); lower strand in blue] was paired with all constructs. The ES2 (arrowhead) and predicted stability (kcal/mol; 3.0 Mfol, M. Zuker) of upstream and downstream duplexes with gRNA are indicated. The ΔG of the downstream duplex for RNAs 1, 2 and 10 is -21.5 , and for RNAs 3–9 is -18.0 . On the left, the pre-mRNA assigned number, size and Mfol predicted stability of the pre-mRNA alone are indicated. (B, D and F) A6 full-round insertion assayed with the indicated peak editing fractions from either Q-Sepharose ('Q') or consecutive Q-Sepharose and DNA-cellulose columns ('D'), and 3' end-labeled pre-mRNA constructs. The pre-mRNA size and assigned number (in parentheses) are indicated [e.g. 72(1) is 72 nt RNA1], as well as accurately edited +3U RNA (open arrowhead). Inaccurate +1U RNA is also indicated when it is the major product (filled arrowhead). Dots indicate sites of spontaneous RNA breakage, typically at U positions. All editing products were confirmed in reactions +/- gRNA (data not shown). (B and D) Constructs with serial truncations are shown, and (F) functional derivatives of (the otherwise inactive) 34 nt RNA8 are shown. (C, E and G) Plots of relative abundance of +3U RNA for all tested constructs normalized to the efficient 45 nt RNA4 substrate in Q (C) or D (E and G) peak fractions, respectively. Standard deviation of two parallel experiments are indicated, as well as the mean (continuous line).

We next decided to compare the effect of A6 pre-mRNA minimization between insertion and deletion. In nature, insertion or deletion occur at separate ESs. Instead of analyzing different ESs, we considered the possibility of using a single pre-mRNA site to directly compare full-round insertion and deletion, i.e. a site in a natural or artificially converted form (i.e. from deletion to insertion or *vice versa*). Previous reports have shown that A6 and CYb ESs can be converted at least at the cleavage step (19,26), and substrates for partial 'pre-cleaved' editing efficiently support the second and third enzymatic steps (15,24), suggesting that relatively simple ES determinants may govern a complete editing cycle. Indeed, A6 (insertion to deletion) conversions have been used to establish multi-site full-round editing substrates (J. Cruz-Reyes, A. Zhelonkina and B. Sollner-Webb, unpublished data), but opposite conversions (deletion to insertion) had not been possible. Conversion of natural ESs in both directions would enable direct comparisons of full-round insertion and deletion 5' of a common pre-mRNA/gRNA downstream 'anchor' duplex. To this end, we introduced three uridylates at ES2 in the 72 nt RNA1 (generating 75 nt RNA11, Figure 2A), as well as appropriate gRNA changes to preserve single-strandedness of the targeted uridylates and other residues between the duplexes, as well as weak pairing of proximal pairs in the upstream duplex (10). Both RNA11 and parental RNA1 form the same upstream and downstream duplexes with gRNA, so that the two substrates differ only in the structure of the internal loop (Figure 2B shows the sequence changes). Figure 2C shows that the converted RNA11 supported accurate 3U deletion, thus enabling analysis of the same truncations in RNAs 2, 3 and 4 (compare with RNAs 12, 13 and 14; Figures 1A and 2A). Notably, deletion at both converted and natural sites was optimal in presence of mM concentrations of an adenine ribonucleotide (19) (data not shown).

We found that the above mutations similarly affected deletion in both Q and D fractions (Figure 2C–E). Interestingly, the 75 nt RNA11 was quite efficient for ES2 deletion, and an upstream 10 nt resection (65 nt RNA12) had a limited effect. In contrast, a downstream 17 nt resection (58 nt RNA13) inhibited deletion by ~50%, and both truncations combined (48 nt RNA14) were strongly inhibitory. Notably, the predicted folding stability of the latter construct is nearly half of the 65 nt RNA12, and about one-third of the 75 nt RNA11. Furthermore, terminal extensions of the inactive 48 nt RNA14 with (either 5' or 3') artificial pyrimidine-rich sequences (65 nt RNA15 or 65 nt RNA16) re-established efficient deletion (Figure 2F). Control reactions devoid of gRNA confirmed the position of the deletion products (data not shown). Dots indicate sites of spontaneous breakage typically at U positions.

Together, the data in Figures 1 and 2 indicate that in both Q and D fractions, the optimal A6 substrate for ES2 insertion was about 20 nt shorter than for ES2 deletion (45 nt RNA4 and 65 nt RNA12, respectively, in Figures 1B and 2B). Strikingly, while the 45 nt RNA4 is the most efficient insertion substrate, the comparable 48 nt RNA14 failed to support deletion. Furthermore, our analysis suggests that distal A6 pre-mRNA composition (beyond the sequence of the shortest RNA analyzed) and folding stability may influence, but are not critical for, efficient ES2 insertion and deletion.

CYb insertion and deletion editing

To determine whether the observed size differences in minimal substrates for A6 deletion and insertion are specific to this pre-mRNA transcript, or are shared in other substrates, we analyzed the CYb model system. To this end, we started with a 66 nt pre-mRNA/gRNA pair for insertion at ES1, originally described by Igo *et al.* [RNA17 in Figure 3A; (11)]. This substrate contains the entire editing domain (including 13 ESs), which only supports insertion *in vivo*. In both Q and D fractions, a 14 nt upstream resection (52 nt RNA18) and a 21 nt resection (45 nt RNA19) caused a gradual decrease in insertion (Figure 3). In contrast, a 45 nt template for A6 insertion (RNA4) was optimal in both Q and D fractions (Figure 1). In further 5' truncations, the 43 nt RNA20 is similarly efficient as the 45 nt RNA19, whereas the 40 nt RNA21 and 37 nt RNA22 were significantly inhibited in both Q and D fractions (Figure 3C and D). The shortest substrate appeared inactive in the D fraction, and only ~10% active in the Q fraction. All plotted quantifications are the average of two independent experiments, each one including replica series (see material and methods). All pre-mRNA constructs tested have similar predicted folding stability (Figure 3A). Thus, we observed efficient CYb insertion in the Q and D fractions with 52 nt and 66 nt substrates (RNA18 and RNA17, respectively). In contrast with CYb insertion, A6 insertion was quite efficient with a shorter 45 nt substrate (RNA4) but relatively inhibited with larger substrates (Figure 1B and C). This difference between the two substrates suggests that editing complexes are sensitive to transcript-specific features, in addition to absolute substrate size.

In vivo CYb editing only involves insertion; however, it is unknown whether or not the CYb transcript is intrinsically resistant to deletion. To address this question, we attempted artificial conversion of the 66 nt RNA17 by introducing uridylates at ES1 (generating 70 nt RNA23, Figure 4A) and corresponding gRNA sequence changes to mimic a deletion site (Figure 4B details the changes). Notably, this construct supported accurate gRNA-directed 3U deletion at the converted ES1 (Figure 4C, lane 1); therefore, the CYb substrate is not intrinsically resistant to deletion editing. This CYb deletion construct allowed us to test the upstream resections (Figure 4A) previously analyzed for CYb insertion. In the Q and D fractions, 14 and 21 nt resections (56 nt RNA24 and 49 nt RNA25, respectively) have relatively small effects. In further 5' truncations, a 44 nt RNA26, 41 nt RNA27 and 36 nt RNA28 were less efficiently processed in both protein preparations. The shortest construct did not exhibit visible editing in the D fraction (Figure 4E, lane 8). Thus, CYb deletion is quite efficient with a ~49 nt substrate on average in both Q and D fractions. In contrast, A6 deletion was significantly inhibited with the equivalent 48 nt RNA14 in both Q and D fractions. Interestingly, CYb insertion and deletion exhibited similar profiles of editing activity relative to substrate size (compare plots in Figures 3 and 4). This could be a reflection of the natural substrate not being exposed to selection pressures associated with deletion editing. Thus, the optimal substrate for deletion differs between A6 and CYb. Such differences in deletion, together with the insertion differences described above, further suggest that pre-mRNA size and nucleotide composition affect editing efficiency in a transcript-specific manner.

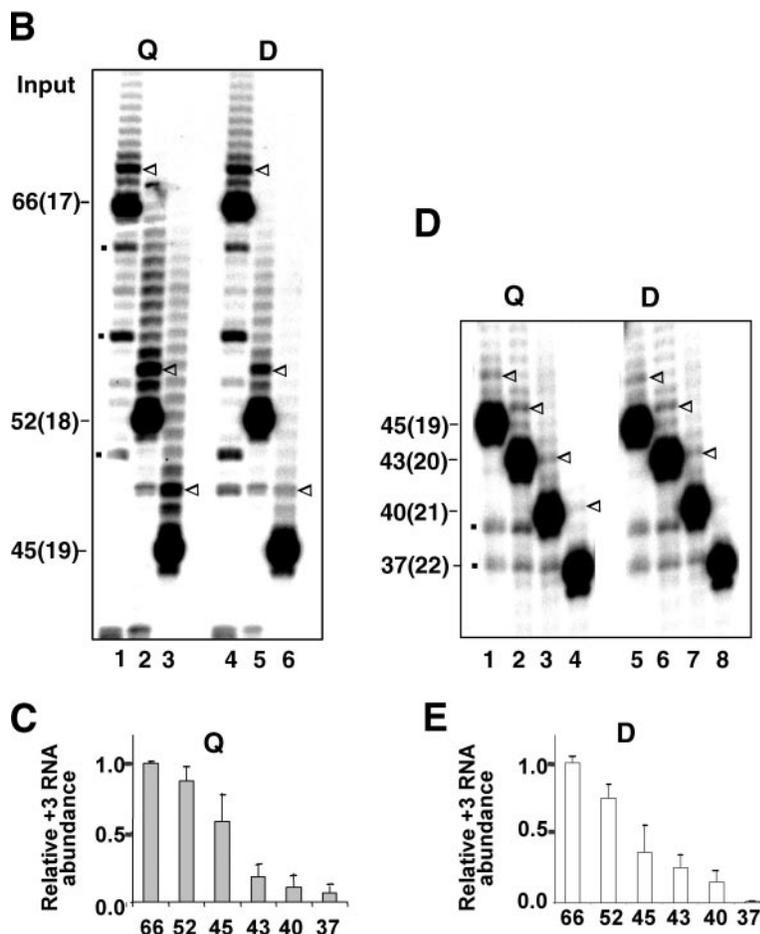
A

Figure 3. Mutagenesis of CYb pre-mRNA for full-round insertion at ES1. (A) Sequence of the starting 66 nt substrate (RNA17), originally described by Igo *et al.* (11) and derivatives. The CYb gRNA is a 3U insertion version of the original gCYb[558]USD-2A (11). All labeling is as in Figure 1. (B and D) CYb insertion assays in Q and D fractions, with accurate 3U insertion products indicated by an arrowhead. (C and E) Plots of relative accumulation of +3U RNA.

to direct 3U deletion at the natural ES2. Interestingly, the largest substrate was most efficient, despite its significant predicted folding stability (Figure 5A), and the shorter 70 nt RNA30 and 55 nt RNA (data not shown) exhibited corresponding decreasing efficiency in both Q and D fractions (Figure 5C) (data not shown).

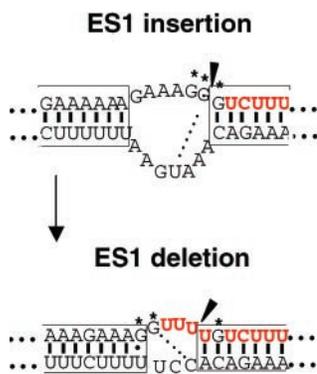
So far we showed conversion from insertion to deletion. To determine if the opposite conversion can be performed (i.e. from deletion to insertion), we altered the natural deletion ES2

in RPS12 to mimic an insertion site, similar to the natural CYb ES1 and A6 ES2 (Figure 5B). Based on our A6 and CYb analysis, we tested a 43 nt construct with a complementary gRNA design directing 3U insertion. This RNA pair forms a potential ligation 'bridge' with residues adjoining the ES, which stimulates insertion but is dispensable for deletion (10,12,15). Also, it favors single-strand character of residues between the duplexes, and enables weak pairing at the 3' end of the upstream duplex, as in deletion sites (10,15). As predicted,

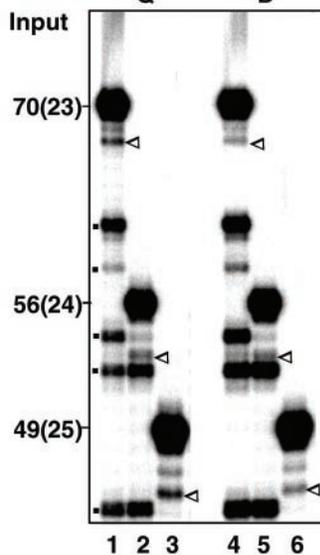
A

Pre- mRNA No.	nt.	ΔG	CYb DELETION
23	70	-10.5	G U U A A G A A U A A U G G U U A U A A A U U U U A U A U A A A A G C G G A G A A A A A A G A A A G G U U U U U G U C U U U U A A U G U C A G C C U C U U U U U U C U U U U U C C U C A C A G A A A A U U A C A G U C A G G G U C C
24	56	-8.2	G A A U A A A U U U U A U A U A A A A G C G G A G A A A A A A G A A A G G U U U U U G U C U U U U A A U G U C A G
25	49	-9.1	G A A A A U A U A A A A G C G G A G A A A A A A G A A A G G U U U U U G U C U U U U A A U G U C A G
26	44	-8.4	G A G A A A A G C G G A G A A A A A A G A A A G G U U U U U G U C U U U U A A U G U C A G
27	41	-7.9	G A A G C G G A G A A A A A A G A A A G G U U U U U G U C U U U U A A U G U C A G
28	36	-7.9	G G G A G A A A A A A A G A A A G G U U U U U G U C U U U U A A U G U C A G

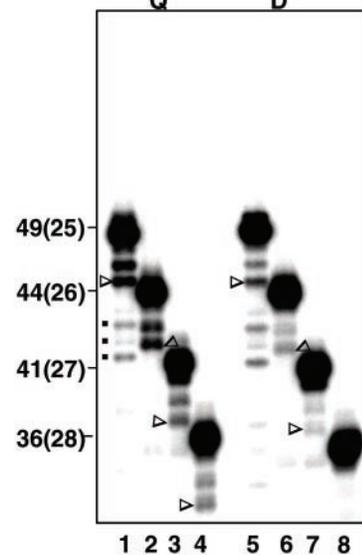
B



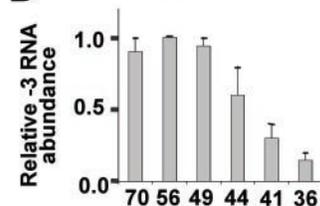
C



E



D



F

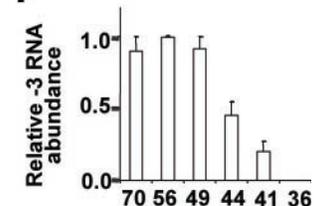


Figure 4. Mutagenesis of CYb pre-mRNA for full-round deletion at converted ES1. (A) Sequence of the starting 70 nt RNA23, which is similar to 66 nt RNA17 in Figure 3, but with converted ES1 for 3U deletion, and derivatives. The gRNA was modified to mimic a guiding region directing deletion. All labeling is as in Figure 1. (B) Diagram of parental and converted ES1 for full-round editing and flanking duplexes. Additional sequence not shown is indicated by the dots. (C and E) CYb deletion assays. (D and F) Plots of relative accumulation of -3U RNA.

the resulting 43 nt RNA31 supported accurate 3U insertion in both Q and D fractions (Figure 5D) (data not shown). Longer 51 nt and shorter 40 nt RNAs appeared slightly more and less efficient, respectively (data not shown). The absolute percent of correctly edited product with the novel RPS12 substrates reported here varied between experiments and protein preparations, but generally ranged between 5 and 10% of remaining pre-mRNA input.

Thus, RPS12 substrates were established for full-round deletion and insertion. Furthermore, the conversion of a natural deletion site into an insertion site confirms that all genetic

information that commit editing complexes to carry out full-round deletion or insertion resides within the ES and suggests that such information can be manipulated to reprogram the mode of the editing machinery at virtually any natural site.

DISCUSSION

In this study, we defined minimal pre-mRNAs substrates for full-round insertion and deletion at specific ESs, and systematically assessed the contribution of distal regions on editing efficiency. We compared editing with complexes moderately

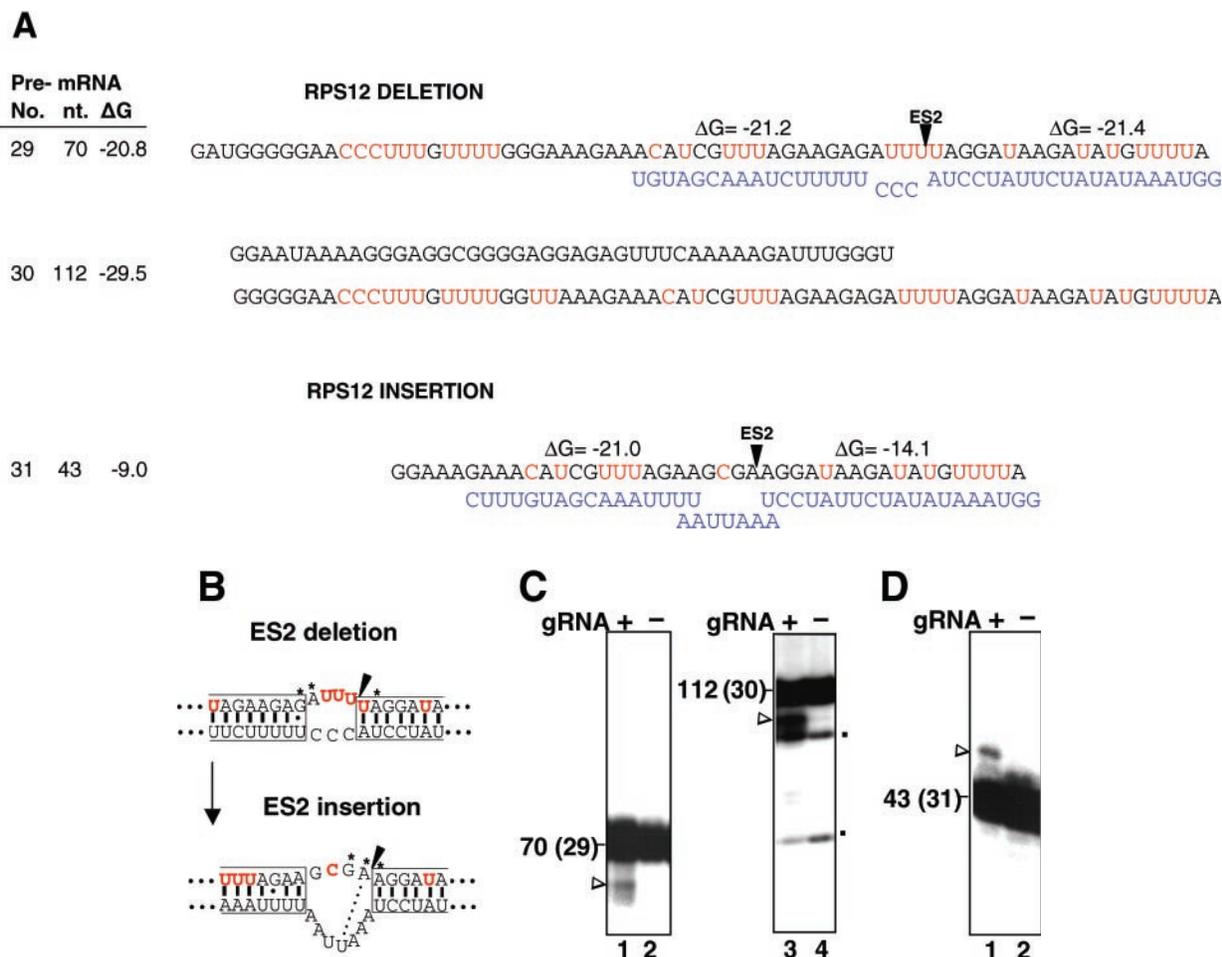


Figure 5. The establishment of an RPS12 pre-mRNA substrate for full-round deletion and insertion at natural and converted ES2. (A) Sequences of 70 nt RNA29 and 112 nt RNA30 for deletion and 43 nt RNA31 for insertion with partially complementary gRNAs. All labeling is as in Figure 1. (B) Diagram of parental deletion and converted insertion ES2. (C and D) Deletion and insertion assays in peak D fraction with controls devoid of gRNA. Additional sequence not shown is indicated by the dots.

enriched by Q-Sepharose (peak editing fraction, 'Q fraction') or significantly enriched by consecutive Q-Sepharose and DNA-cellulose (peak fraction, 'D fraction') columns (28). Three different model transcript substrates were analyzed, A6, CYb and RPS12 (the latter established in this study). Interestingly, in some reactions, substrate shortening caused greater inhibition of insertion activity in the D fraction than in the Q fraction (e.g. see Figures 1D and 4E), suggesting that stringent purification of editing complexes causes partial loss or inactivation of a relevant RNA-binding activity. Furthermore, at least in the case of A6, substrate shortening negatively affected the fidelity of insertion (Figure 1D, compare lanes 1 and 5, or 6 and 10). Our proposed loss of a relevant factor upon extensive purification of editing complexes may explain why some protein preparations by other laboratories support efficient pre-cleaved editing (using short substrates), but not full-round editing reactions.

The best substrate size identified for each transcript analyzed was similar in both Q and D peak fractions (e.g. A6 insertion was best with a ~ 45 nt substrate in both protein preparations); however, significant differences between deletion and insertion were observed: ~ 45 and ~ 65 nt A6

substrates were best for insertion and deletion, respectively (Figures 1 and 2), and longer substrates were somewhat less efficient in deletion. Notably, in contrast to our optimized ~ 45 nt insertion substrate, the corresponding 48 nt deletion substrate (with 3Us at ES2 accounting for the size difference) was strongly inhibited (Figure 2D and E).

In the case of CYb editing, insertion was most efficient with a 66 nt transcript (Figure 3) (15), whereas deletion was best with 49 and 56 nt substrates (Figure 4). Significantly larger CYb insertion substrates (~ 80 and 100 nt long) are somewhat less efficient (data not shown). The molecular basis of these differences between deletion and insertion in the same substrate, and between the two substrates, is presently unclear. They may reflect the complexity of substrate recognition due to the proposed partition of functions associated with insertion and deletion in editing complexes, and the evidently large diversity of nucleotide sequences proximal and distal to sites encountered by editing complexes. Thus, it is feasible that one or more relevant RNA-protein recognitions differ between full-round insertion and deletion. Furthermore, such interactions appear to be significantly influenced by transcript-specific features in and around editing sites.

Purine-rich regions typically found upstream of editing sites have been proposed to anneal with natural gRNA 3' (poly)U tails (4,7), and structural studies in trypanosomes have confirmed this interaction (14). Whether or not purine-rich regions or other pre-mRNA features exhibit additional *cis*-acting functions has not been determined in trypanosomes. In *Leishmania*, ND7 and CYb pre-mRNA substrates that are devoid of natural purine-rich regions are functional, although the presence of an uncharacterized small 5' determinant near the ND7 editing site was proposed (17,29). Furthermore, a separate 34 nt A/U element, in CYb pre-mRNA, appears to modulate gRNA-directed, and induces gRNA-independent, insertion (30). In our current study, inactivating resections of the A6 upstream purine-rich region were rescued by pyrimidine-rich extensions. Similarly, inactivating downstream resections can be rescued by heterologous sequences. This indicates that the specific sequence and purine-richness of the truncated A6 regions have little if any modulatory functions. Whether more proximal 5' and 3' sequences (not dissected in this study) bear specific determinants remains to be analyzed. Our *in vitro* analysis of CYb has not revealed yet the presence of critical *cis*-elements in substrates up to 100 nt long (this work and data not shown), but further studies may be necessary.

Based on our observations with the A6 and CYb systems, we established RPS12 substrates for deletion and converted insertion at ES2. Interestingly, an ~100 nt long RSP12 substrate supports the best deletion level thus far, and shorter substrates (70 nt and 55 nt long; Figure 5C and data not shown) are correspondingly less efficient. Similar to the A6 system, we obtained efficient insertion with shorter substrates (~43–50 nt long, Figure 5C) (data not shown) than those needed for optimal deletion. We suspect that efficient model substrates of transcripts naturally undergoing both types of editing and deletion may be shorter for insertion than for deletion. CYb may be an exception as this transcript has not been exposed to natural pressures imposed by deletion editing.

Finally, the effect of pre-mRNA minimization on full-round deletion and insertion was directly compared at natural and artificially converted sites in three different substrates. Such conversions were performed in both directions: from deletion to insertion and from insertion to deletion, thus confirming expectations that the type of editing is exclusively determined by mRNA and gRNA residues within the ES (19,26). Importantly, both natural and converted sites for full-round deletion studied here exhibited the reported [mM] adenine ribonucleotide requirement for optimal cleavage. In contrast, both natural and converted insertion sites do not require such [mM] adenine ribonucleotide supplement (19). Thus, both the type of editing and associated nucleotide requirements were converted. Sites for partial (pre-cleaved) deletion and insertion editing have also been established on significantly altered pre-mRNA sequences (15,24). Site conversion within the CYb editing domain appears surprising as this substrate is only known to undergo insertion *in vivo*. If the CYb editing domain is not intrinsically refractory to U deletion, inserted or genomically encoded Us may be occasionally removed by 'misguiding' (31), or by proofreading cycles (to eliminate extra U insertions) with cognate gRNAs. It is also conceivable that the natural substrate (1151 bases) uses specialized features, and/or CYb-specific binding factors to downregulate deletion functions of bound complexes.

ACKNOWLEDGEMENTS

We thank Drs. Linda Guarino, Andy C. LiWang and C.T. Ranjith-Kumar for comments on the manuscript and helpful discussions. This work was supported by a grant from the NIH (GM067130) to JC-R. Funding to pay the Open Access publication charges for this article was provided by the same grant above.

Conflict of interest statement. None declared.

REFERENCES

- Madison-Antenucci, S., Grams, J. and Hajduk, S.L. (2002) Editing machines: the complexities of trypanosome RNA editing. *Cell*, **108**, 435–438.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- Blom, D., Burg, J.V., 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.
- Muller, U.F., Lambert, L. and Goringer, H.U. (2001) Annealing of RNA editing substrates facilitated by guide RNA-binding protein gBP21. *EMBO J.*, **20**, 1394–1404.
- Cruz-Reyes, J., Rusche, L.N., Piller, K.J. and Sollner-Webb, B. (2001) Trypanosome RNA editing: simple guide RNA features enhance U deletion 100-fold. *Mol. Cell Biol.*, **21**, 884–892.
- 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.
- 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. (1999) Mapping contacts between gRNA and mRNA in trypanosome RNA editing. *Nucleic Acids Res.*, **27**, 778–787.
- Leung, S.S. and Koslowsky, D.J. (2001) RNA editing in *Trypanosoma brucei*: characterization of gRNA U-tail interactions with partially edited mRNA substrates. *Nucleic Acids Res.*, **29**, 703–709.
- Igo, R.P. Jr, Palazzo, S.S., Burgess, M.L., Panigrahi, A.K. and Stuart, K. (2000) Uridylyate addition and RNA ligation contribute to the specificity of kinetoplastid insertion RNA editing. *Mol. Cell Biol.*, **20**, 8447–8457.
- Kapushoc, S.T. and Simpson, L. (1999) *In vitro* uridine insertion RNA editing mediated by *cis*-acting guide RNAs. *RNA*, **5**, 656–669.
- 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.
- Seiwert, S.D. and Stuart, K. (1994) RNA editing: transfer of genetic information from gRNA to precursor mRNA *in vitro*. *Science*, **266**, 114–117.

19. 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.
20. 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.
21. 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.
22. 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.
23. 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.
24. Igo, R.P.Jr, Weston, D.S., Ernst, N.L., Panigrahi, A.K., Salavati, R. and Stuart, K. (2002) Role of uridylyate-specific exoribonuclease activity in *Trypanosoma brucei* RNA editing. *Eukaryot Cell*, **1**, 112–118.
25. 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.
26. Sollner-Webb, B., Rusche, L.N. and Cruz-Reyes, J. (2001) Ribonuclease activities of trypanosome RNA editing complex directed to cleave specifically at a chosen site. *Methods Enzymol.*, **341**, 154–174.
27. Harris, M.E. and Hajduk, S.L. (1992) Kinetoplastid RNA editing: in vitro formation of cytochrome *b* gRNA–mRNA chimeras from synthetic substrate RNAs. *Cell*, **68**, 1091–1099.
28. 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.
29. 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.
30. Opegard, L.M., Kabb, A.L. and Connell, G.J. (2000) Activation of guide RNA-directed editing of a cytochrome *b* mRNA. *J. Biol. Chem.*, **275**, 33911–33919.
31. 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.