6944–6965 *Nucleic Acids Research, 2023, Vol. 51, No. 13* *Published online 29 May 2023*
https://doi.org/10.1093/nar/gkad453

Trypanosome RNA helicase KREH2 differentially controls non-canonical editing and putative repressive structure via a novel proposed 'bifunctional' gRNA in mRNA A6

Joshua Meehan[1], Suzanne M. McDermott[2,3], Alasdair Ivens[3], Zachary Goodall[1], Zihao Chen[4], Zihao Yu[1], Jia Woo[5], Tyler Rodshagen[2], Laura McCleskey[1], Rebecca Sechrist[1], Kenneth Stuart[2,3], Lanying Zeng[1], Silvi Rouskin[5], Nicholas J. Savill[4], Achim Schnaufer[4], Xiuren Zhang[1] and Jorge Cruz-Reyes[1,*]

[1]Department of Biochemistry and Biophysics, Texas A&M University, College Station, TX 77843, USA, [2]Center for Global Infectious Disease Research, Seattle Children's Research Institute, Seattle, WA 98109, USA, [3]Departments of Pediatrics and Global Health, University of Washington School of Medicine, Seattle, WA, USA, [4]Institute of Immunology and Infection Research, University of Edinburgh, Edinburgh EH9 3FL, UK and [5]Department of Microbiology, Harvard Medical School, Boston, MA 02115, USA

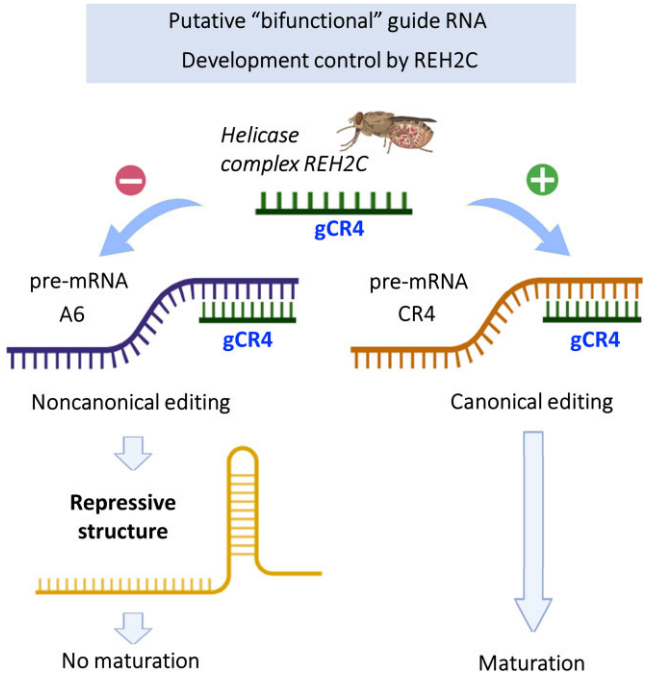
Received December 21, 2022; Revised April 07, 2023; Editorial Decision April 27, 2023; Accepted May 26, 2023

ABSTRACT

U-insertion/deletion (U-indel) RNA editing in trypanosome mitochondria is directed by guide RNAs (gRNAs). This editing may developmentally control respiration in bloodstream forms (BSF) and insect procyclic forms (PCF). Holo-editosomes include the accessory RNA Editing Substrate Binding Complex (RESC) and RNA Editing Helicase 2 Complex (REH2C), but the specific proteins controlling differential editing remain unknown. Also, RNA editing appears highly error prone because most U-indels do not match the canonical pattern. However, despite extensive non-canonical editing of unknown functions, accurate canonical editing is required for normal cell growth. In PCF, REH2C controls editing fidelity in RESC-bound mRNAs. Here, we report that KREH2, a REH2C-associated helicase, developmentally controls programmed non-canonical editing, including an abundant 3′ element in ATPase subunit 6 (A6) mRNA. The 3′ element sequence is directed by a proposed novel regulatory gRNA. In PCF, KREH2 RNAi-knockdown up-regulates the 3′ element, which establishes a stable structure hindering element removal by canonical initiator-gRNA-directed editing. In BSF, KREH2-knockdown does not up-regulate the 3′ element but reduces its high abundance. Thus, KREH2 differentially controls extensive non-canonical editing and associated RNA structure via a novel regulatory gRNA, potentially hijacking factors as a 'molecular sponge'. Furthermore, this gRNA is bifunctional, serving in canonical CR4 mRNA editing whilst installing a structural element in A6 mRNA.

*To whom correspondence should be addressed. Tel: +1 979 324 8249; Fax: +1 979 324 8249; Email: cruzrey@tamu.edu

GRAPHICAL ABSTRACT



INTRODUCTION

Trypanosoma brucei is a member of the protist group Euglenozoa and causes human African trypanosomiasis (1–3). This parasite has a life cycle that traverses between the tsetse fly vector and the mammalian host, where it proliferates as procyclic and bloodstream forms (PCF and BSF), respectively. Trypanosomes exhibit unique genetic and biological phenomena, including mitochondrial RNA (mtRNA) editing through site-specific insertion and deletion of uridines (U-indels). Twelve of the 18 primary mRNA transcripts lack an open reading frame (ORF), which has to be established post-transcriptionally via precise U-indels. This editing may control respiratory physiology and is developmentally regulated. However, the regulatory editing factors and their *modus operandi* during the life cycle remain to be uncovered.

The mitochondrial genome (kinetoplast or kDNA) is a planar network of catenated maxicircle and minicircle relaxed molecules in trypanosomes and related trypanosomatids (4,5). In *T. brucei*, maxicircles contain genes encoding rRNA, ribosomal protein S12 (RPS12), ATPase subunit 6 (A6) and proteins in respiratory complexes. Most mitochondrial mRNAs require extensive editing, which can double the primary transcript size. Other mRNAs require moderate editing or are never edited (6). Minicircles encode ∼1000 different guide RNAs (gRNAs; ∼45–60 nt) that exhibit complementarity to reported canonically edited sequences. However, gRNAs only show combined pairings and mismatches with pre-edited mRNA. Typical gRNAs in *T. brucei* include an anchor region that initiates binding with mRNA through Watson–Crick pairing to form a partial duplex and a guiding region that directs the U-indels by Watson–Crick and GU wobble pairing. The average anchor and guiding regions are 6–11 nt and 20–40 nt long, respectively (7,8). gRNAs usually also carry encoded 5'- and 3'-terminal bases that are not used in canonical editing and a post-transcriptionally added 3'-oligo(U) tail. A few gRNAs carry 5'-oligo(A) tails that could be added post-transcriptionally or via RNA polymerase transcription slippage events (9). Editing progresses from 3' to 5' in overlapping blocks, each block directed by a gRNA (7,10). Thus, the canonical ORF is installed based on complementarity by canonical gRNAs. However, only a few molecules match the canonical pattern at steady state, while most carry 'incorrect' non-canonical edits. Non-canonical U-indels usually appear in 'editing junctions' of variable length and composition representing regions of ongoing editing, which are flanked by 3'-canonical and 5'-pre-edited sequences (11,12). The 3′ ends of gRNAs are sometimes found ligated to a targeted editing site *in vivo*. Such 'bimolecular chimeras' can occur during the basic editing reaction but are most probably not true intermediates. However, chimeras are diagnostic of on-target gRNA pairing with cognate mRNA *in vivo* (13,14).

During development, *T. brucei* requires massive changes in metabolism, including an energetic switch in its single mitochondrion (15–17). PCFs employ cytochrome-mediated oxidative phosphorylation. However, BSF parasites lack cytochromes and some Krebs cycle enzymes, and produce ATP by glycolysis since sugar is plentiful in serum. Parasite adaptation to different host environments includes massive remodeling of the edited transcriptome. Canonical 'fully edited' sequences that encode cytochrome mRNAs (complexes III and IV) readily accumulate in PCFs but are barely detectable (or absent) in BSFs. Other mRNAs, e.g. for complex I (NADH dehydrogenase), exhibit significant differences in editing between the two stages, and some transcripts, e.g. subunit 6 (A6) of the F_1F_O-ATPase complex (complex V) or RPS12, are thought to be similarly edited in both stages. Pre-mRNA and gRNA levels are relatively constant at steady state, so these organisms may not regulate transcript availability but rather maturation, including editing (9,10). Thus, U-indel editing is essential and may modulate mitochondrial function during the parasite life cycle. However, the key editing regulatory proteins and specific molecular mechanisms under their control during development remain unknown.

The editosome holoenzyme is a dynamic supramolecular ribonucleoprotein structure of variable composition and organization. Holo-editosomes include ∼40 nuclear-encoded proteins in three multiprotein complexes and additional factors: RNA Editing Catalytic Complex (RECC), RNA Editing Substrate Binding Complex (RESC) and RNA Editing Helicase 2 Complex (REH2C) (11,18,19). RESC appears to be a platform for mRNA–gRNA hybrid assembly and action by RECC and REH2C (20–24). RECC, the first identified editing complex, has three specialized isoforms that catalyze endonucleolytic cleavage, U-indels and ligation (25–28). RESC is heterogeneous and contains gRNAs (29–32). REH2C includes three core proteins, DExH-Box RNA helicase KREH2, KH2F1 and KH2F2, and has ATP-dependent 3′–5′ double-stranded RNA (dsRNA) unwinding activity. KREH2 is the scaffold for KH2F1 and KH2F2 cofactor association since each cofactor co-purifies with KREH2 upon depletion of the other. PCF knockdown of KREH2 or KH2F1 inhibited cell growth and

editing (21,33–35). RNA interference (RNAi) of KREH2 or KH2F1 similarly decreased total editing and editing fidelity in the examined transcripts (21,35–37). We and others initially showed enrichment of editing substrates and products in purified RESC (20,22) and showed that REH2C affects editing of RESC-bound mRNAs *in trans* (21,35,37).

U-indel editing is considered highly error prone and possibly even energetically 'wasteful' (38); however, basic molecular mechanisms must ensure enough accuracy within this process to provide mature mitochondrial transcripts needed to support cell growth. Eukaryotes use different strategies to regulate the fidelity of RNA processing events, including RNA helicases that interrogate protein–RNA interactions and remodel RNA structure (39). Amidst high-frequency editing that does not match the canonical pattern, the mechanisms that regulate fidelity in U-indel editing need to be defined. Our prior amplicon-RNA-seq studies in RESC-bound and total mtRNA in PCF cells showed that REH2C controls editing fidelity in mRNA RPS12 and mRNA A6 (henceforth A6) in a site-specific and mRNA-specific fashion (37) (Table 1). REH2C loss of function via RNAi knockdown of KREH2 or KH2F1 affected fidelity preferentially by inhibiting canonical editing at sites in the 5' half of RPS12. However, RNAi knockdown of KH2F1 affected fidelity by increasing non-canonical editing along the A6 fragment examined, but preferentially at 3′ sites. These KH2F1 knockdown-induced non-canonical edits formed an abundant alternative 3′ sequence that appeared to block canonical A6 editing. However, we did not characterize the responsible repression mechanism. Significantly, KH2F1 stabilizes KREH2 in PCF cells, so KH2F1-RNAi silenced both proteins simultaneously in this life cycle stage (21,37). Thus, the resulting dual knockdown left unclear whether specific depletion of KREH2 would induce the formation of the non-canonical 3' element in PCF, BSF cells or both.

Besides the abundant A6 3' element in PCFs mentioned above, alternative non-canonical editing has previously been described in *T. brucei* and *Leptomonas pyrrhocoris* (9,20,40–42), where it could potentially impact coding capacity (43) or derail canonical editing (44). Non-canonical gRNAs that could direct this alternative editing were identified in some cases, but whether specific editing factors may regulate their use was not defined. Also, specific bases outside the guiding region may direct relevant non-canonical edits. In RPS12, an abundant non-canonical 2U-insertion event, which pauses 3′–5′ progression, particularly in RESC-bound transcripts, may be directed by two conserved 3'-terminal adenines in the initiator gRNA-1 (37).

Here, we report the first example of a protein factor, KREH2, that differentially controls non-canonical editing in A6 mRNA in different ways: firstly, by introducing a specialized 3′ sequence; and secondly, by affecting non-canonical editing along A6 and so 3′–5′ progression. The specialized 3′-non-canonical editing is abundant, particularly in RESC-bound transcripts. A6 maturation is constitutive in PCF and BSF cells; however, KREH2 depletion differently disrupts non-canonical editing in the two stages. We propose a PCF-specific model of A6 3′ editing control upon knockdown of KREH2, including two main steps. First, KREH2-RNAi induces up-regulation of 3′-non-canonical editing by a novel gRNA type. This gRNA installs an alternative 3′-high-frequency element (3′-HFE) and abutting pre-edited sequence at the first few A6 sites. Together, these two sequences create an extended 3′ element that fully complements the putative regulatory gRNA. Second, the extended 3′ element forms a stable structure that sequesters the A6 3' terminus, occluding this region from potential 'repair' by canonical initiator gRNA-1-directed editing. We applied DMS-MaPseq technology (45) to analyze this RNA structure experimentally.

This A6 extended 3′ element was relatively more abundant in BSF than in PCF cells, but knockdown of KREH2 did not up-regulate its formation in BSFs. Furthermore, the identified novel gRNA type may be bifunctional, exhibiting opposing dual roles, i.e. positive, promoting canonical editing progression in CR4 mRNA versus negative, installing a non-canonical structural element in A6. *In vivo* detection of bimolecular chimeras between the proposed bifunctional gRNA and mRNAs A6 or CR4 also indicates on-target contacts. These studies support a general model whereby not all non-canonical editing is metabolically wasteful, challenging the view of editing as an example of constructive neutral evolution. Because of the programmed (genome-encoded and regulated) non-canonical editing reported here, the helicase complex, REH2C, potentially controls novel regulatory gRNAs and repressive RNA structure to modulate the production of proteins and overall mitochondrial physiology.

MATERIALS AND METHODS

PCF and BSF cell culture and transfection

Trypanosoma brucei strain Lister 427 29-13 PCF cells were grown in selective SDM-79 medium supplemented with 10% tetracycline (Tet)-tested fetal bovine serum (FBS) (R&D Systems). Cell lines were maintained in log phase growth with regular dilutions to maintain a density of $< 1 \times 10^7$ cells/ml. Cell lines were selected with the following concentrations of antibiotics: 15 μg/ml G418 sulfate (Thermo Fisher), 50 μg/ml hygromycin (Invivogen), 2.5 μg/ml phleomycin (Thermo Fisher) and 1 μg/ml puromycin (Invivogen), when applicable. Growth curves were carried out in 24-well plates and counted on a Beckman-Coulter Counter Z2. Log phase cell cultures were seeded at 2×10^6 cells/ml in biological replicates and counted every other day, diluting back to 2×10^6 cells/ml after every count.

Trypanosoma brucei strain Lister 427 BSF cells were grown in selective HMI-9 medium supplemented with 10% Tet-tested FBS. Cell lines were maintained in log phase growth with daily dilutions to maintain a density of $< 1 \times 10^6$ cells/ml. Cell lines were selected with the following concentrations of antibiotics: 1 μg/ml blasticidin (Thermo Fisher), 2.5 μg/ml G418 sulfate (Thermo Fisher), 2.5 μg/ml phleomycin (Thermo Fisher) and 0.1 μg/ml puromycin (Invivogen), when applicable. All RNAi constructs were made as described previously and induced with 1 μg/ml Tet (Sigma) (35,37).

Transfections were carried out in an Amaxa Nucleofector 2b Device (Lonza) with 2 mm gap electroporation

Table 1. Glossary of terms

Sites	Any position between two non-T nucleotides (cDNA) in the reference T-stripped sequence. Sites are numbered 3′ to 5′ in the direction of editing. Editing events just 5′ to a G, C or A are scored.
T number	The number of T nucleotides (cDNA) immediately 5′ to a G, C or A. T numbers between 0 and 16 were scored.
Canonical editing site (ES)	Any position between two non-T nucleotides (in cDNA) where the T number in the canonical pattern (mature) is expected. Other sites are not modified in mature mRNA.
Pre-edited (PE) sequence	Transcript sequence which has the genomic encoded T number at all sites.
Fully-edited sequence	Transcript sequence which contains the exact T number at all sites in the canonical pattern.
Canonical guide RNA (gRNA)	A gRNA which directs a block of editing events that match the published canonical pattern.
Alternative non-canonical gRNA	A gRNA which directs a block of editing events that differ from the published canonical pattern.
Canonical (C) value	The total number of reads at each site with the expected T number in the canonical pattern.
Non-canonical (NC) value	The total number of reads at each site with a T number that differs from the canonical pattern.
Total editing value	The total number of reads at each site that contain any T number, except for the T number in the pre-edited sequence.
NC/C ratio	A normalized value at each position that scores overall deviation from the expected canonical pattern 'editing fidelity' matching gRNAs classified as canonical. This value is determined by dividing the NC value by the C value at each site.
Fold change in NC/C ratio	The relative change in NC/C ratio between two consecutive sites, 3′ to 5′. Instances where the fold change is significant suggest intrinsic pause sites. Large (>5-fold) fold changes indicate major pausing sites (MPSs) in canonical editing progression.
KREH2-RNAi enhanced pause sites	An intrinsic PS, including MPS, in which the immediate 5′ site (exhibiting high NC/C ratio) shows further decreased fidelity (i.e. even higher NC/C) upon KREH2-RNAi.
3′-high-frequency element (3′-HFE) in A6 mRNA	An abundant non-canonical sequence revealed by very high NC/C values across block 1 in A6 mRNA. This 3′-HFE is installed by a novel regulatory anti-initiator gRNA (below) and helps prevent canonical A6 editing.
Repressive RNA fold in A6 mRNA	A stable secondary structure determined by DMS-MapSeq. This fold made by an ∼42 nt element (including the 3′-HFE) may block all canonical A6 editing.
Bifunctional gRNA	A gRNA with a putative dual function, e.g. canonical editing progression in mRNA CR4, and non-canonical as a putative repressor (anti-initiator) blocking canonical gRNA-1 in HFE-containing mRNA A6.

cuvettes (VWR) using the nucleofector program X-001. For transfection, ∼1 × 10^7 PCF or BSF cells were pelleted at 2500 *g* at 4°C for 10 min. Cell pellets were resuspended in 200 μl of transfection buffer (90 mM sodium phosphate, 5 mM potassium chloride, 0.15 mM calcium chloride, 50 mM HEPES, pH 7.3), transferred to 2 mm gap electroporation cuvettes (BTX) and immediately electroporated. Cells were then added to 1 ml of selection-free HMI-9 medium with 20% FBS and diluted in 24-well culture plates at 1:50, 1:200 and 1:1000 dilutions alongside a negative control plate (no DNA electroporation). After ∼16 h recovery time, cells were diluted using medium with selection and maintained until control plates no longer supported growth. RNAi cell lines were verified as previously described for PCF KREH2 RNAi using quantitative polymerase chain reaction (qPCR) and western blots (35,37). Growth curves were carried out in 24-well plates and counted on a Beckman-Coulter Counter Z2. Log phase cell cultures were seeded at 5 × 10^4 cells/ml in biological replicates and counted daily, diluting back to 5 × 10^4 cells/ml daily.

Real-time qPCR analysis

Total RNA was harvested from 2 × 10^8 cells using TRIzol following RNAi induction by addition of 1 μg/ml Tet for 3 days (BSF) or 4 days (PCF). Isolated RNA was treated with 10 U of Turbo DNase (Life Technologies) according to the manufacturer's instructions, and purified through acid phenol:chloroform extraction. A 2 μg aliquot of total RNA was used to synthesize cDNA using the iScript select cDNA synthesis kit with random hexamers. cDNAs were then pre-amplified in multiplex specific-target-amplification (STA) reactions using TaqMan PreAmp master mix (Life Technologies) and with the following thermocycling conditions: 1 cycle at 95°C for 10 min and 14 cycles of 95°C for 15 s and 60°C for 4 min. Pre-amplified cDNA was treated with exonuclease I (New England Biolabs) and diluted 5- (for BSF) or 10-fold (for PCF). High-throughput real-time PCR was then conducted on the BioMark HD system with Fluidigm 48-by-48 dynamic array integrated fluidic circuits (IFCs), using SsoFast EvaGreen supermix with Low ROX (Bio-Rad) and primers described in Supplementary Table S1. Processing of the IFCs and operation of the instruments were performed according to the manufacturer's procedures. PCR was performed using the thermal protocol GE Fast 96 × 96 PCR + Melt (v2.pcl). Data were analyzed in the Fluidigm real-time PCR analysis software, using the linear (derivative) baseline correction method and the auto (global) threshold cycle (CT) method. The CT values determined were exported to Excel software for further processing. Calculations of fold changes in RNA levels in samples following RNAi induction, relative to no induction, were done using the 2 [−ΔΔC(T)] method (46) using TERT as an internal reference (47). Technical quadruplicates of each cDNA sample were assayed for each target and internal reference per experiment and C(T) data averaged before

performing the 2 [−ΔΔC(T)] calculation. Experiments were repeated using three biological replicates, as in prior studies (48).

Immunofluorescence of *T. brucei* editing proteins

Immunofluorescent microscopy of *T. brucei* cells was carried out using 10 ml of BSF cells at ≤ 1 × 10^6 cells/ml. Cells were pelleted at 2500 *g*, 4°C for 10 min then washed with 1 ml of IF wash buffer [1 mM $CaCl_2$, 1 mM $MgCl_2$, 1× phosphate-buffered saline (PBS) pH 7.4] and pelleted at 2500 *g*, 4°C for 10 min. Cells were resuspended in fixing solution (4% paraformaldehyde in 1× PBS pH 7.4) to give a final concentration of 1 × 10^7 cells/ml and incubated at room temperature for 5 min. Fixed cells were applied to a poly-L-lysine-coated 4-well chambered coverglass (Thermo Fisher) The slide had been previously treated with 1 M KOH for 1 h at room temperature, rinsed three times with distilled water and allowed to air dry. Poly-L-lysine coating was done at 37°C and allowed to dry by laminar flow. A 150 μl aliquot of fixed cells (∼1.5 × 10^6 cells) was pipetted into chambers of the slide and allowed to adhere for 10 min at room temperature. The supernatant was removed by pipetting along the side of the chamber, and chambers were washed by gently pipetting 250 μl of IF wash buffer, incubating for 3 min then removing and replacing the buffer with a pipette three times. Cells were permeabilized with IF permeabilization buffer (0.1% Triton X-100 in 1× PBS pH 7.4) for 10 min at room temperature before washing with three changes of IF wash buffer. Cells were then blocked with 250 μl of blocking buffer [4% bovine serum albumin (BSA) in 1× PBS pH 7.4] for 30 min at room temperature. Blocking buffer was removed by pipetting, and 250 μl of primary antibody, diluted in blocking buffer, was individually added to chambers and incubated at room temperature for 45 min. Dilutions for antibodies were 1:1000 KH2F1, 1:5000 RESC1/2 and 1:1 KREL1 (i.e. 125 μl of blocking buffer and 125 μl of of antibody). Slides were then washed with three changes of IF wash buffer and stained with 10 μg/ml 4′,6-diamidino-2-phenylindole (DAPI) for 10 min then washed with three changes of IF wash buffer. Cells were treated with 1:750 diluted anti-rabbit IgG Atto 488 (Sigma-Aldrich) and 1:375 AlexaFluor 647 donkey anti-mouse IgG (Invitrogen) secondary antibody for 1 h at room temperature. Slides were washed with three changes of IF wash buffer and dried with an air gun. Slides were imaged on a Nikon Eclipse Ti inverted epifluorescence microscope using a ×100 objective (Plan Apochromat, NA 1.45, oil immersion) and images analyzed with NIS-Elements Advanced Research software packages.

Mitochondria-enriched extracts for Illumina and protein analyses

RNAi was induced with Tet for 4 days in PCF cells and for 3 days in BSF cells. These time points were determined by growth curves and western blotting of editing proteins. Total mtRNA was prepared as described previously (37) with slight modifications. For Illumina analyses, mitochondria-enriched extracts were prepared from four independent replicate flasks. For PCFs, 100 ml of cell culture was grown per replicate to a density of 1.3 × 10^7–1.7 × 10^7 cells/ml. For BSFs, 150 ml of cell culture was grown per replicate to a density of 0.8 × 10^6–1.3 × 10^6 cells/ml. Cells were pelleted, and mitochondrial vesicles were extracted by dissolving cell pellets in 500 μl of DTE buffer (1 mM Tris, pH 8.0, 1 mM EDTA) followed by six strokes of dounce in a glass tissue homogenizer with a tight-fitting pestle. The homogenized lysate was treated with 10 U of Turbo DNase I at 4°C for 1 h, and vesicles were pelleted at 16 000 *g* for 10 min. The resulting mitochondria-enriched pellet was lysed with the addition of 200 μl of 1× MRB lysis buffer [25 mM Tris, pH 8.0, 10 mM magnesium acetate, 60 mM KCl, 5 mM dithiothreitol (DTT), 1 mM EDTA, 0.5% Triton X-100, 10% glycerol] supplemented with 1× Roche cOmplete protease inhibitor cocktail (Sigma) and incubation on ice for 20 min. Mitochondrial debris was pelleted at 18 000 *g* and the resulting supernatant was stored at –80°C until used for sodium dodecylsulfate–polyacrylamide gel electrophoresis (SDS–PAGE) or immunoprecipitations. Western blots of protein subunits in REH2C and RESC complexes in SDS–PAGE were performed as previously described (21,34).

Preparation of RNA for library construction

RNA was isolated from four biological replicates of BSF and PCF cells ± KREH2-RNAi induction. For total mtRNA samples, mitochondrial vesicles were enriched from 4.5 × 10^8 PCF cells or 1.6 × 10^8 BSF cells. The mitochondrial pellet was resuspended in 1 ml of TRIzol reagent and RNA isolated as per the manufacturer's instructions. A 10 μg aliquot of isolated RNA was DNase treated using 50 U of DNase I (Thermo Fisher) in the presence of 10 U of Superase (Thermo) and purified once again through acid phenol:chloroform extraction before cDNA synthesis.

RESC6 immunoprecipitations were also performed as described with slight modifications (21). Protein A Dynabeads (Invitrogen) were blocked with 4% BSA (Millipore) in 1× MRB wash buffer (25 mM Tris, pH 8.0, 150 mM NaCl, 1 mM EDTA, 10 mM magnesium acetate) for 1 h at 4°C with shaking at 700 rpm. Antibodies were conjugated to beads at a ratio of 1:1 (beads:antibodies) in 1× MRB washing buffer overnight at 4°C with shaking at 700 rpm. Immunoprecipitations were performed for technical replicates, and the resulting RNA was combined for each sample. For each immunoprecipitation, 1.5 × 10^8 cell equivalents of PCF mitochondrial extract was added to 0.6 mg of conjugated beads. Proteins were bound to beads by incubating at 4°C for 90 min with shaking at 700 rpm, and occasionally flicking to prevent the beads from fully settling. Beads were separated on a magnetic rack, and the supernatant was discarded. Bead pellets were resuspended in 1 ml of 1× MRB wash buffer and separated on a magnetic rack for five washes, changing buffer each time. The bead pellet was then resuspended to give a final bead concentration of 30 mg/ml. For each sample, 1.2 mg of recovered beads were then treated with 100 μl of protein digestion buffer containing 8 U of proteinase K (NEB) and 0.5% SDS for 2 h at 50°C. RNA was then isolated using acid phenol:chloroform extraction (Sigma) according to the manufacturer's di-

rections. Isolated RNA was precipitated in isopropanol at –80°C overnight and resuspended in 20 μl of DNase solution containing 50 U of DNase I and 10 U of Superase, and incubated at 37°C for 1 h. RNA was purified again by acid phenol:chloroform extraction and ethanol precipitated at –80°C overnight before resuspending in 10 μl of diethylpyrocarbonate (DEPC)-treated H_2O.

cDNA synthesis, Illumina sample preparation and sequencing

A6-specific cDNA synthesis was carried out using an oligonucleotide that anchors to the never-edited 3′ region of A6 (primer 2607) (Supplementary Table S1: synthetic DNA and constructs). cDNA synthesis was carried out with 2 μg of mtRNA using the iScript Select cDNA Synthesis Kit (BioRad). We checked for the specificity of targeted cDNA synthesis and subsequent amplification as follows. BSF and PCF gene-specific cDNAs were amplified using KAPA Hifi HotStart ReadyMix (Roche) utilizing oligos containing universal Illumina adapters (1683/1684) and purified using the Nucleospin PCR clean-up kit (TakaraBio). Purified PCR products were then amplified with oligos 2542 and 2543 to generate amplicons with terminal 5′ HindIII and 3′ BamHI sites. These amplicons and pLEW100v5 plasmid were digested with BamHI-HF (NEB) and HindIII-HF (NEB), and purified in 1% agarose gels using the NucleoSpin Gel and PCR Clean-up Kit. Digested amplicon and pLEW100v5 plasmid were ligated using T4 DNA ligase (NEB) at a 5:1 insert:vector molar ratio to produce plasmid p599. Ligated plasmid was transformed into Stellar competent cells (TakaraBio). Ten individual bacterial colonies were picked and Sanger sequenced using oligos 2542/2543 for full coverage of the amplicon with forward and reverse primers. Illumina libraries were prepared as previously described (37) with modifications. A6 libraries were amplified from 10 ng of BSF or PCF gene-specific cDNA for 24 cycles with oligos containing Illumina adapters (1683/1684). Amplicons containing adapters were purified using Ampure XP PCR purification beads (Beckman-Coulter) and visualized on a 4200 Agilent Tapestation. A 10 ng aliquot of each amplicon was indexed using the Nextera XT Index Kit V2 SetA (Illumina) with 12 cycles of amplification. The resultant Illumina libraries were then purified again with Ampure XP beads, and their concentration was measured using a QuBit 4 Fluorometer and high-sensitivity dsDNA QuBit Reagent (Thermo Fisher). Libraries were diluted to 4 nM and pooled. The library pool was denatured with 0.1 M NaOH and diluted to 8 pM as per the manufacturer's instructions in HT1 buffer with 40% PhiX spike-in. Samples were run with an Illumina MiSeq Reagent Nano Kit v2 2 × 250, which produced 1.56×10^6 paired reads after QC filtering and removal of PhiX spike-in. A total of 3.16×10^5 reads were identified as putative A6 transcripts by their A|C|G nucleotide content.

Processing RNA-seq data in A6 editing and identification of non-canonical gRNA isoforms that complement the extended 3′ element

Amplicon RNA-seq of A6 editing was processed as reported (37). Subsequently, sample alignment output data were further processed in the R environment (http://www.r-project.org) for summarizing and figure generation purposes. Searches for gRNAs encoded in the *T. brucei* EATRO 1125 minicircle genome (7) that match alternatively edited mRNA sequences were performed using Python scripts (package 3.7) as previously described (8). Alignments of predicted (annotated in minicircles) and sequenced gRNA in EATRO 1125 total mtRNA (7,8) are available online at http://hank.bio.ed.ac.uk. Alignments of sequenced gRNA in EATRO 164 total mtRNA (10) and Lister 427 in RESC6 immunoprecipitations and total mtRNA (20) are available online at http://bioserv.mps.ohio-state.edu/RNAseq/T-brucei/MRBs/.

Calculations and statistical analysis

Total editing and NC/C (non-canonical/canonical) ratio values (previously termed Inc/Cor ratio) of the percentage of reads, both site-by-site and cumulative, were calculated as previously reported (37). In this study, NC (non-canonical) and C (canonical) replaced the former Inc (incorrect) and Cor (correct) terms, respectively. This updated nomenclature reflects a key observation in the current study that specialized non-canonical editing is controlled by editing proteins. Graphs compare replicate sets for two different conditions (e.g. BSF versus PCF, or –Tet versus +Tet), where one replicate set includes at least three biological replicates, and another set includes at least two biological replicates. These replicate sets enabled statistical calculation of *P*-values, average and standard deviation (SD). A description of samples and *P*-values for all sets compared are included (Supplementary Table S2). Fold change values were calculated as reported (Supplementary Table S3) (37). A limited amount of KH2F1-RNAi data in PCF cells examined the frequency of the 3′ element and the most 5′ position in cumulative plots. To generate *P*-values for the effects of KH2F1-RNAi, we combined reported A6 data for 3 and 4 days of RNAi ($n = 2$ each day) since the 2 days of treatment had comparable outcomes. This allowed us to increase the total number of replicates for the +Tet condition. The KH2F1 data statistics shown here were not previously reported. We used one-way analysis of variance (ANOVA) to test the null hypothesis that there is no significant difference between groups, with this null hypothesis rejected at $P < 0.05$. The mean ± SD of independent biological replicates was reported.

DMS-MaPseq for experimental determination of RNA structure

The RNA structure of full-length A6 was experimentally determined *in vitro* by DMS-MaPseq (49). Synthetic gBlocks (IDT DNA Technologies) were generated containing the entire A6 pre-edited (PE, gBlock 2637) or the most common isoform bearing the 3′-extended element described in this study (HFE, gBlock 2638) (Supplementary Table S1). These sequences were amplified with oligos 2631 and 2632 to produce T7-coupled amplicons with a 5′ HindIII and 3′ BamHI site. PCR product and pHD1344Tub(PAC) plasmid (a gift from Suzanne McDermott) was digested with HindIII-HF and BamHI-HF, and gel purified. Purified products were ligated with T4 DNA ligase at a 5:1

molar ratio and transformed into Stellar competent cells. Five bacterial colonies were picked for each construct (p614 PE and p615 HFE), and colony PCR was performed using the internal oligo 2631 and external oligo 2242, and verified by agarose gel. A 15 ml aliquot of culture from positive colonies was grown overnight in selective LB medium (100 μg/ml ampicillin) in a 37°C incubator with shaking at 200 rpm. Plasmid DNA was isolated and sent for Sanger sequencing in technical replicates using the forward oligo 2622 and reverse oligo 2611 to confirm the design of the constructs.

Once confirmed, 10 μg of plasmid was linearized using XhoI (NEB) at 37°C overnight. Plasmid digestion was confirmed by agarose gel. Linearized plasmid was cleaned up using phenol–chloroform extraction, and 1 μg of linearized plasmid was used for run-off transcription using the T7 Megascript kit (Thermo Fisher) at 37°C for 6 h. Template plasmid was digested by adding 5 U of Turbo DNase and incubated at 37°C for 15 min. Synthesized RNA was isolated using the Zymo Cleanup kit (Zymo Research). A 2 μg aliquot of purified RNA in 10 μl was denatured at 95°C for 1 min. Then 89.5 μl of 1× Refolding Buffer, which consists of 397 mM sodium cacodylate buffer (Electron Microscopy Sciences) and 6 mM $MgCl_2$, was added to denatured RNA on ice, and the RNA was allowed to refold at 37°C for 20 min. Dimethyl sulfate (DMS; Fisher Scientific) at 0.5% (v/v) was added, and the reaction was incubated at 37°C for 4 min with shaking at 800 rpm. The reaction was quenched with the addition of 60 μl of 2-mercaptoethanol, and RNA was cleaned up using the Zymo RNA cleanup kit.

DMS-MaPseq library generation and reactivity analysis

DMS-MaPseq libraries were generated using IDT's xGenTM Broad-Range RNA Library Prep Kit with slight modifications. A 500 ng of purified DMS-modified RNA was used as input. Briefly, RNA was fragmented for 2 min according to the manufacturer's instructions without adding reagent F2 (dNTPs). After 2 min, the fragmentation mix was placed on ice immediately. The mixture was then incubated with the reverse transcription mix [1 μl of TGIRT (Ingex), 1 μl of water, 1 μl of enzyme R1 (RNase inhibitor) and 1 μl of DTT] at room temperature for 30 min. Then, F2 (dNTPs) was added, and the fragmented RNA mixture was reverse transcribed under the conditions: 20°C for 10 min, 42°C for 10 min, 55°C for 60 min and denaturation by adding 1 μl of 4 M NaOH at 95°C for 3 min. To neutralize the mixture, 2 μl of 4 M HCl was added, and the volume of this mixture was brought up to 50 μl with nuclease-free water. Then, the reverse-transcribed cDNA was cleaned using a 1× volume ratio of SPRI beads (Beckman-Coulter) and eluted in 10 μl of EDTA TE. Samples were then adapted, extended, ligated and amplified for eight cycles for A6 following IDT's instructions. The libraries (∼300–400 bp) were gel-purified on an 8% TBE polyacrylamide gel (Thermo Fisher) and precipitated using isopropanol. To sequence the libraries, samples were loaded on an iSeq-100 sequencing flow cell with the iSeq-100 High-throughput sequencing kit and the libraries were run on iSeq-100 (paired-end run, 2 × 151 cycles).

FASTQ files were processed and analyzed to determine the DMS signal using the DREEM (Detection of RNA folding Ensembles using Expectation-Maximization clustering) pipeline (49). Briefly, reads were trimmed using TrimGalore (github.com/FelixKrueger/TrimGalore) to remove Illumina adapters. Trimmed paired reads were then mapped to the *T. brucei* EATRO 1125 maxicircle genome (accession: MK584625) (8) using Bowtie2 with the parameters: –loc– –no-un– –no-discorda– –no-mixed -L 12 -X 1000. For each pair of aligned reads, a bit vector was generated and the mutational signatures were analyzed using the DREEM algorithm (49). To quantify the population average DMS reactivity at each position, the ratio of mismatches and deletions to total coverage at each nucleotide position was calculated. DMS reactivities were normalized to the median of the top 5% of DMS reactivities to a scale of 0 to 1. These normalized DMS reactivities were used as folding constraints for predicting RNA secondary structures with the program RNAstructure v.6.0.1 (50). RNA secondary structures were visualized using VARNA v.3.93 (51).

Isolation of *in vivo* chimeric molecules of gRNA gCR4 with mRNAs A6 or CR4

To isolate gRNA/mRNA bimolecular chimeras *in vivo*, we generated A6 or CR4 gene-specific cDNA as described in earlier sections using 2 μg of DNase-treated mtRNA from wild-type PCF cells. cDNA was generated using the BioRad iScript Select cDNA synthesis kit acccording to the manufacturer's directions with oligo 2607 (A6) or 2789 (CR4) in a 20 μl reaction. A 2 μl aliquot of resulting cDNA was used in a PCR to amplify chimeric RNA sequences in a 50 μl Phusion HF polymerase PCR using the forward oligo 2875 (CR4 gRNA) and 2833 (A6 mRNA) or 2877 (CR4 mRNA). These primers are designed to amplify chimeric sequences of CR4 gRNA with either A6 (pre-edited at the first five edit sites) or canonically edited CR4 mRNA; and produce a 5′ HindIII and 3′ XhoI restriction site overhang. The resulting PCR product was verified on a 2% agarose gel and the corresponding bands were gel eluted using the Nucleospin PCR cleanup kit. Undigested PCR product was cloned into HindIII/XhoI-digested pHD plasmid using the In-Fusion cloning kit at a 2:1 molar ratio as described earlier for *in vitro* DMS mapping. This reaction generated plasmids p652 (CR4-A6) and p653 (CR4-CR4). Then 2 μl of In-Fusion product was transformed into Stellar competent cells according to the manufacturer's instructions, plated on ampicillin-selective LB agar plates and allowed to grow overnight at 37°C. Ten individual colonies were picked from each plate (CR4-A6 or CR4-CR4), and plasmid was isolated from cultures using a QIAgen mini plasmid isolation kit. A 25 ng aliquot of isolated plasmid was used as template in a 50 μl Phusion HF polymerase PCR with oligos 2056/2242 to amplify a portion of the vector containing our cloned fragment. These PCR amplicons were then purified by 1% agarose gel electrophoresis and 150 ng of PCR product was analyzed by Sanger sequencing in technical replicates by MCLAB using oligo 2611 and 2622. Alignments of chimera sequences were done using the MUSCLE alignment tool.

RESULTS

KREH2 is required for efficient editing maturation of a broad range of substrates examined in BSF and PCF *T. brucei*

To examine the importance of KREH2 in BSF cells for the first time, we generated a Tet-regulatable KREH2-RNAi cell line in this life cycle stage. KREH2-RNAi reduced KREH2 protein levels by day two post-induction, and induced a growth defect by day four post-induction (Figure 1A). However, other editing proteins examined, including RESC1 and RESC2, were not affected during these first few days of RNAi (Figure 1B), suggesting an absence of secondary effects during this period. All subsequent studies in BSF cells were performed after 3 days of RNAi. We previously reported this KREH2-RNAi construct in PCF cells, which targets the 3′-untranslated region (UTR), and we also examined these cells after 3 days of RNAi (21,35).

To begin, we compared the overall effect of KREH2-RNAi on editing in PCF and BSF cells by performing RT-qPCR on 9 out of the 12 editing targets in mitochondria (Figure 1C). The target mRNAs examined included: pan edited A6, RPS12, CO3, ND3, ND7 (both 5′ and 3′ domains) and ND8; and minimally edited CYb, CO2 and MURF2. These assays measure canonically edited or pre-edited sequences at the 5′ or 3′ end of the editing domain, respectively, in each target so that the scored amplicons represent fully edited products or pre-edited substrates. Our analyses showed that KREH2 depletion reduces fully edited pan-edited and minimally edited targets at steady state in BSF and PCF cells. Pan-edited substrates appeared more affected than minimally edited substrates, except for MURF2. However, changes in fully edited ND7 in PCF cells seemed less evident than in other edited targets. Similar RT-qPCR results in PCF cells were previously reported with a different KREH2-RNAi construct which targets the ORF (33).

We noted that large decreases in fully edited transcripts following KREH2 knockdown do not cause corresponding increases in their pre-edited precursors. This discrepancy has also been reported in RNAi studies of other editing proteins (24,33). It may reflect high stability or elevated levels of pre-edited mRNAs in cells or changes in partially edited intermediates, which RT-qPCR does not measure. KREH2 knockdown did not affect mitochondrial transcripts examined that do not undergo editing (ND4 mRNA, 9S rRNA and 12S rRNA). In immunofluorescence analyses in BSF cells, we showed that representative proteins in RESC and RECC, RESC1/2 and KREL1, respectively, localize near the kDNA. Previous reports showed similar localization of editing proteins in PCFs (33,52). We also showed for the first time that KH2F1 in REH2C localizes near kDNA in BSFs (Figure 1D). The above observations indicate that normal KREH2 expression is necessary for cell growth and canonical RNA editing in BSF and PCF cells.

KREH2 knockdown differentially affects total editing at the 3’ end of A6 in BSF and PCF cells

Our above assays confirmed that KREH2 is necessary for growth and editing in PCF *T. brucei* (33,34) and, for the first time, for growth and editing of BSF *T. brucei*. However, these assays do not inform on the effects of KREH2 down-regulation on 3′–5’ editing progression. We applied base-resolution RNA-Seq of an A6 3′ fragment, including ORF (78 sites) and UTR (22 sites) sequences (Supplementary Figure S1A), to examine early editing progression upon knockdown of KREH2 in PCF and BSF cells. We began by examining A6 in total mtRNA. To illustrate the raw data initially collected and tallied by our bioinformatics pipeline, we included a stack plot of a representative biological replicate in wild-type PCF and BSF cells (Figure 2A, B). These plots provide a snapshot of all editing events scored at each site in the A6 amplicon. We focused on the 3’ fragment of A6 because this region exhibits extensive editing action guided by the first four canonical gRNAs (∼30% of the sequence examined). The first gRNA, initiator gRNA-1, and the second gRNA, gRNA-2, cover most of the 3’-UTR in A6, which may offer early checkpoints to regulate the entire canonical editing cascade. Following this distinctive 3’ region, the remaining A6 sequence examined (almost two-thirds of the amplicon) exhibited far fewer total edits. This profile of editing action along A6, first reported in PCF cells (37), is similarly found in BSF cells (Supplementary Figure S1B–D).

In the amplified A6 3′ fragment, the forward primer matches a pre-edited sequence, and the reverse primer matches a never-edited 3’ sequence. The forward primer may select for amplicons with low editing action at the 5’ end. However, because editing action drops dramatically before the 5′ half of the amplicon is reached, a high editing action at the 3’ region seems intrinsic to A6 in both PCF and BSF cells. The initiator gRNA-1 in our alignments is one of two reported potential alternative initiator gRNAs in A6 canonical editing (9). We reported that alternative gRNA-1 gA6 (774–822) in strain EATRO164 (alias isoform gA6 *B1.alt* in strain Lister 427) produced the best match with edited A6 examined by Sanger (20) and Illumina sequencing (37) in the PCF strain Lister 427. Recent studies of A6 in strain Lister by another lab have also used gA6 (774–822) in strain EATRO164 (alias gA6 *B1.alt* in Lister 427) (44). We did not find a match between the other possible candidate initiator gRNA and A6 edited molecules in our samples. gRNA-2 in these and our prior A6 studies in PCF cells was originally annotated in minicircle DNA libraries in strain EATRO1125 (8). These gRNA-1 and gRNA-2 sequences generate the best match with the canonical A6 pattern in our BSF samples.

The A6 editing profile is generally similar in PCF and BSF cells. However, the two stages exhibit evident differences in partial non-canonical edits (yellow bars), particularly across the first two gRNAs in A6, where we found the most editing action. We focused on the 3’ region to examine the effects of REH2C loss of function on early editing progression. All subsequent analyses directly compared independent biological replicates of each sample plus or minus KREH2 knockdown. Samples without knockdown showed marked differences between PCF and BSF cells in analyses of total editing action at the A6 3′ terminus across the initiator gRNA-1. In particular, sites 31–38 exhibited significantly higher total editing in BSF versus PCF cells (Figure 2C; Supplementary Figure S2), although plots of cumulative total editing at the 5′-most site along the

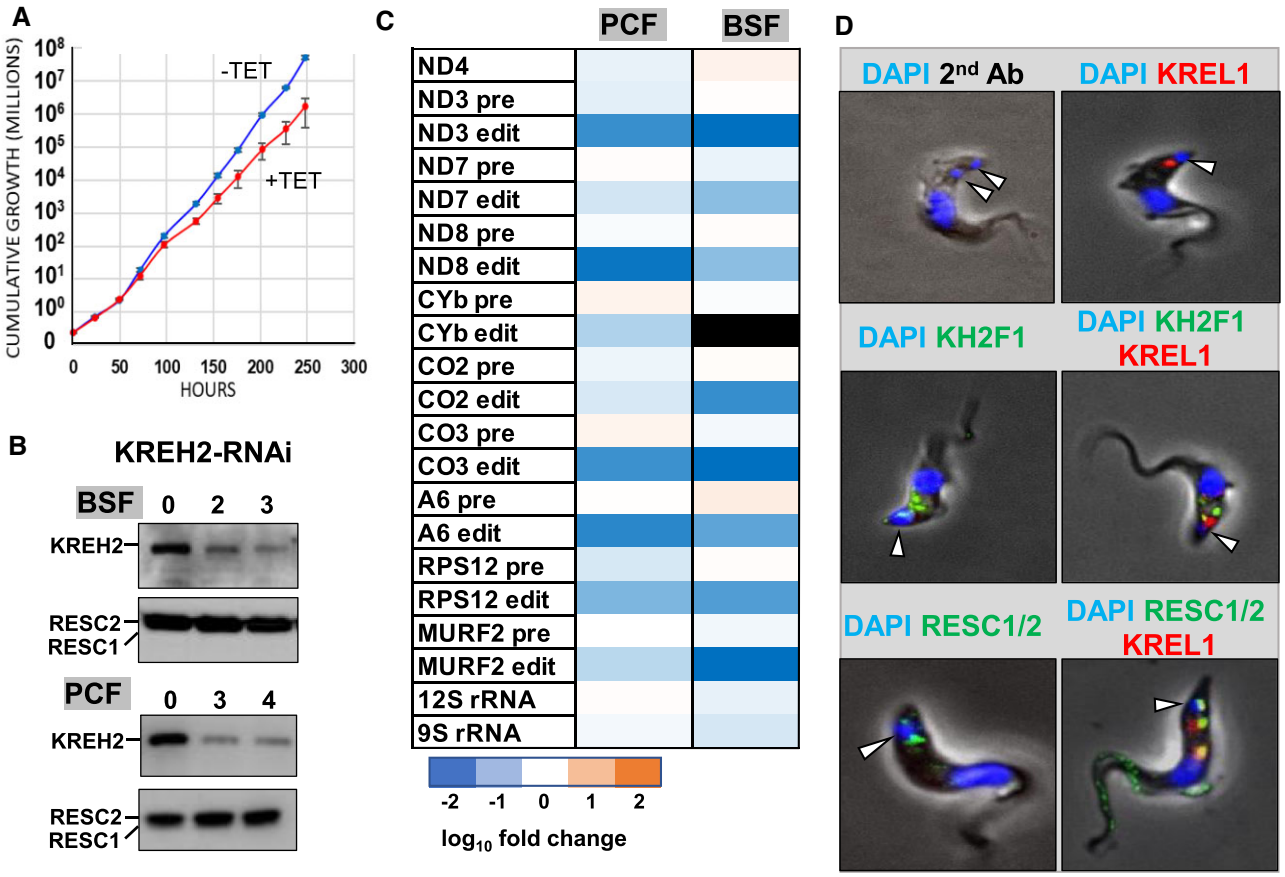


Figure 1. KREH2 knockdown effect on a panel of editing mRNAs in PCF and BSF cells. (**A**) Growth curve of KREH2-RNAi ± Tet in PCF and BSF cells. (**B**) Western blot of PCF and BSF mitochondrial extracts with KREH2-RNAi ± Tet at the indicated days post-induction. KREH2 and RESC1/2 were examined. RESC1/2 were used to control for loading and secondary effects, e.g. stability of other extract proteins. (**C**) Heat map of RT-qPCR assays of a panel of fully edited and pre-edited transcripts upon KREH2-RNAi in PCF (left) and BSF (right) total mtRNA. Plotted values represent the log10 relative abundance of –Tet vesus +Tet samples. Assays were normalized to a housekeeping gene (TERT) and the –Tet control. Plotted values are the average of three biological replicates per condition ($n = 3$). Mitochondrial ND4 mRNA and 12S and 9S rRNA transcripts do not undergo editing. (**D**) Immunofluorescent microscopy of BSF *T. brucei* cells. Cells were imaged for marker proteins in editing complexes: KH2F1 (REH2C), RESC1/2 (RESC) or KREL1 (RECC). White arrows point to DAPI-stained kDNA in each cell.

amplicon examined exhibited significantly higher total editing action in PCF than in BSF cells (Figure 2D). Surprisingly, upon KREH2 knockdown, plots of cumulative total editing showed an opposite effect of KREH2-RNAi in PCF versus BSF cells (Figure 2E; Supplementary Figure S2). Namely, cumulative total editing increased in PCFs but decreased in BSFs upon KREH2-RNAi. This opposite effect was unexpected because A6 editing is not thought to be developmentally regulated (2,44). Overall, these results identify the first editing protein, KREH2, that exhibits a differential effect on A6 editing progression in PCF and BSF cells. This differential effect on total edits was observed at most sites examined in A6, with large differences observed preferentially during early editing across the initiator gRNA-1. We also examined the entire A6 amplicon in total editing and other analyses described below (Supplementary Figures S1–S4; Supplementary Tables S2 and S3) and observed the same effects in cumulative counts across the entire amplicon, confirming that the major differential changes had occurred in early editing.

KREH2 knockdown differentially affects relative editing fidelity at the 3' end of A6 in PCF and BSF cells

We wondered whether the observed differential effects of KREH2 knockdown on total editing in PCF and BSF cells reflect changes in relative editing fidelity (i.e. normalized NC/C ratio: the percentage of non-canonical reads over canonical reads at individual sites), particularly in the 3'-UTR in A6, where early editing may be regulated. To address this possibility, we plotted site-by-site and cumulative NC/C values across the entire amplicon while focusing on the first two gRNAs, which cover most of the 3'-UTR in A6 (Figure 3; Supplementary Figure S1A). Large NC/C values reveal substantial editing action which deviates from the expected editing pattern, i.e. they indicate low editing fidelity.

We first compared editing fidelity in PCF versus BSF mtRNA (Figure 3A, B). As expected, based on our snapshots, the stretch spanning editing sites 31–38 in BSF mtRNA included some of the highest NC/C values examined in our A6 samples. This short stretch in BSF mtRNA

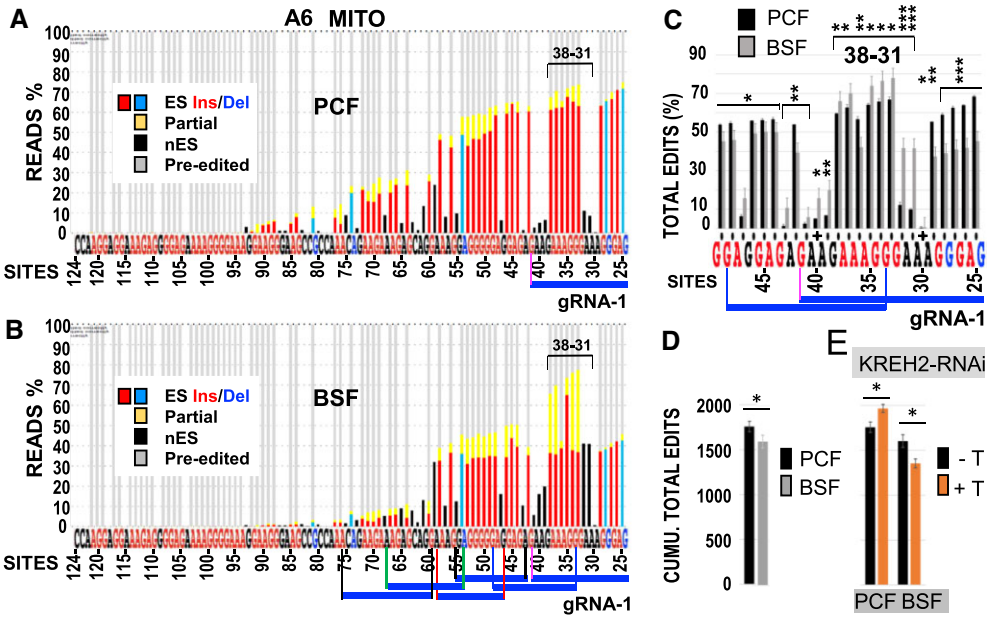


Figure 2. Analyses of A6 total editing in PCF and BSF cells and KREH2-RNAi effects. (**A**) PCF and (**B**) BSF 'snapshots' of typical datasets collected by targeted RNA-seq analyses of amplicons in this study. Stacked histograms show all possible types of editing events at each site in representative replicate samples of mtRNA (Mito) –Tet. Color-coded nucleotides are just 3′ to: canonical sites for U-insertion (Ins, red), U-deletion (Del, blue) or sites not expected to change in mature mRNA (black) (see Table 1; glossary of terms). Bars represent the percentage of canonical insertion (red) or deletion (blue), or non-canonical edits (yellow) at canonical sites, or edits at sites not expected to change (black). Canonical gRNA editing blocks (indigo lines): initiating gRNA-1 or the first few gRNAs (A or B, respectively). (**C**) Site-by-site analysis of total edits across gRNA-1 and gRNA-2 (through site 50). PCF versus BSF mtRNA ± KREH2-RNAi are compared. (**D**) Cumulative total edits in PCF versus BSF mtRNA from uninduced cells. (**E**) Cumulative total edits in PCF and BSF mtRNA ± KREH2-RNAi. The cumulative value at the most 5′ site (site 124) in the amplicon was plotted. Full amplicon analyses are available in Supplementary Figures S2 and S4. Average and error bars of biological replicates ± Tet (± T; $n = 3$) and *P*-values $^{***}P < 0.005$, $^{**}P < 0.05$, $^{*}P < 0.5$ were annotated.

also included dramatic changes in the NC/C ratio between adjacent sites (i.e. NC/C fold change values from one site to the next; see the glossary of terms in Table 1), including at the transitions 30–31 and 36–37. Many instances along A6 where the fold change is significant suggest intrinsic pause sites in canonical editing progression 3′–5'. Large fold change values (>5; arbitrary cut-off) indicate positions of major pausing (37). In such transitions, the 3′ site is referred to as a major pause site (MPS); Supplementary Table S3.

We next asked if KREH2-RNAi similarly affects relative editing fidelity in A6 in PCF and BSF cells, particularly across the first two gRNAs. Surprisingly, this knockdown had the opposite effect on mtRNA in the two stages. Namely, the relative A6 editing fidelity decreased in PCF but increased in BSF cells upon KREH2-RNAi (i.e. cumulative NC/C increased or decreased, respectively; Figure 3B; Supplementary Figure S3C). Most sites examined showed significantly reduced editing accuracy upon KREH2-RNAi in PCF mtRNA (Supplementary Figure S3A). Cumulative ratios at an upstream location (site 70) confirmed that the KREH2-mediated changes are significant (Figure 3C). Differential changes in fidelity by KREH2-RNAi in PCF and BSF mtRNA included sites 31–38 in early editing. A prior study in PCFs showed that RNAi-knockdown of the zinc finger protein KH2F1 in the REH2C complex destabilized KREH2, and decreased A6 editing fidelity (37), as we found here with KREH2-RNAi. Cumulative ratios at an upstream location (site 70) confirmed a significant loss in editing fidelity upon KH2F1-RNAi in PCF mtRNA (Figure 3D). Thus, specific depletion of the RNA helicase KREH2, which does not affect the integrity of KH2F1 (21), decreases the relative editing fidelity along the A6 fragment examined in PCF mtRNA.

We previously reported enrichment of mRNA editing substrates and products in native RESC6 antibody immunoprecipitation versus total mtRNA in PCF cells (20,35,37). From here on, we will refer to RESC-associated mRNA or just 'RESC' to indicate mRNA isolated from native RESC6 immunoprecipitations. Thus, we predicted that KREH2-mediated effects in A6 editing fidelity would be observed at higher frequencies in RESC-associated transcripts. Indeed, site-by-site and cumulative plots in PCF cells showed larger NC/C ratios along the examined A6 sequence in RESC versus mtRNA, including the stretch spanning sites 31–38, in early editing (Figure 3E, F; Supplementary Figures S3A and S4C; compare PCF data in black bars). Cumulative ratios at an upstream location (site 70) confirmed a significant decrease in editing fidelity of A6 in native RESC upon either KREH2-RNAi or KH2F1-RNAi (Figure 3G, H). As mentioned above, besides sites 31–38, KREH2-RNAi significantly affected editing fidelity at other sites, including at intrinsic pause sites, including MPSs, in total mtRNA and RESC (Figure 3F, I; Supplementary Figures S3G and S4G; Supplementary Table S3). We note that KREH2-RNAi-enhanced MPSs are

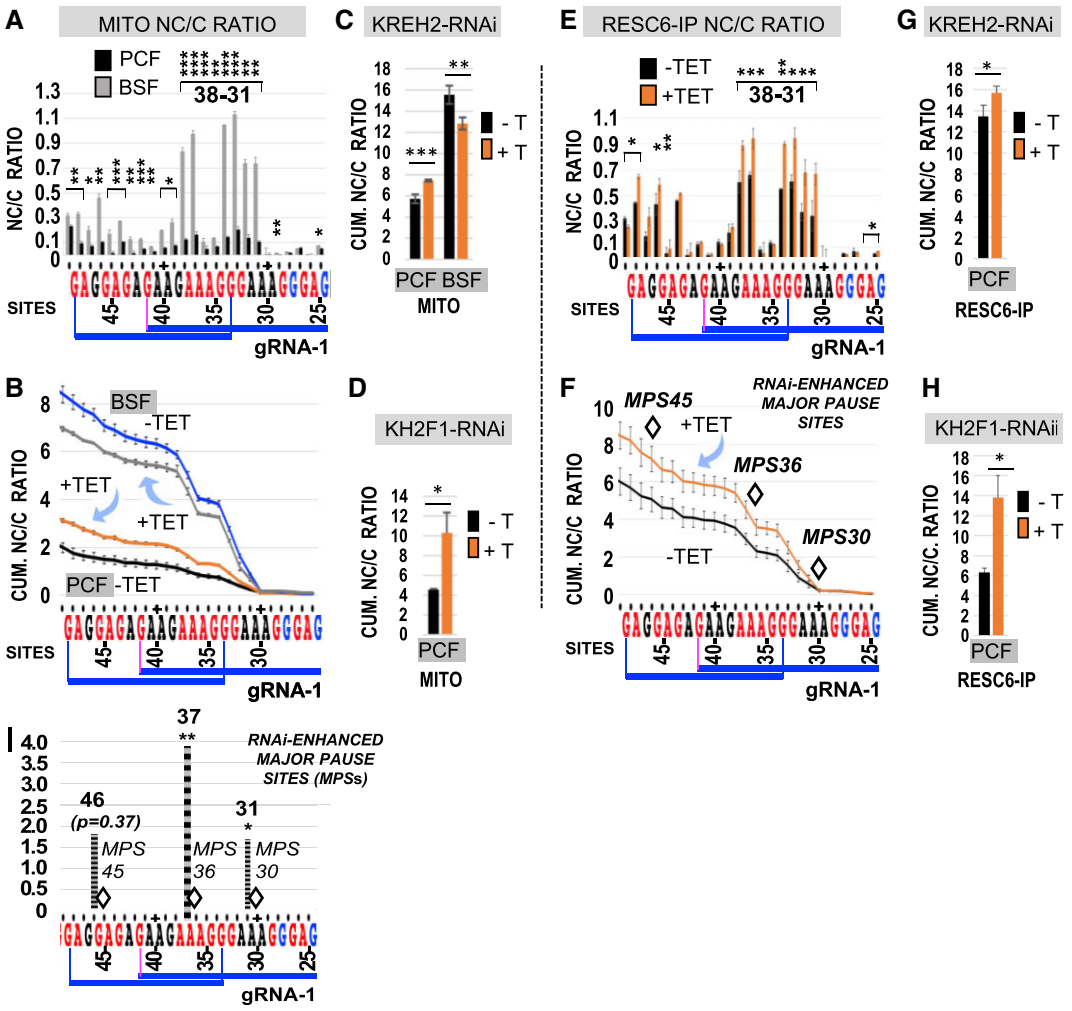


Figure 3. Analyses of A6 NC/C ratios and effect of KREH2-RNAi or KH2F1-RNAi on mtRNA or RESC. (**A**) Site-by-site NC/C ratios in PCF versus BSF mtRNA (Mito) across gRNA blocks 1–2. NC/C ratios are scored as the percentage of non-canonical reads divided by the percentage of canonical reads at the same site. Sites 31–38 (highlighted) exhibited particularly high NC/C ratios. –Tet replicates were used ($n = 3$). (**B**) Cumulative NC/C ratios in PCF versus BSF mtRNA ± KREH2-RNAi. Note the opposite effect of KREH2-RNAi on editing accuracy between the two stages, i.e. editing accuracy decreased in PCF but increased in BSF cells, within the sites examined. (**C**) Cumulative NC/C ratio in PCF versus BSF mtRNA ± KREH2-RNAi. Ratios are through site 70. (**D**) Same as (C) but in PCF mtRNA ± KH2F1-RNAi. (**E**) Site-by-site and (**F**) cumulative NC/C ratio in PCF RESC ± KREH2-RNAi across gRNA blocks 1–2. KREH2-RNAi-enhanced major intrinsic editing pause sites (MPSs) are annotated (diamonds). (**G**) Cumulative NC/C ratio in PCF RESC ± KREH2-RNAi through site 70. (**H**) The same as (G) but in PCF RESC ± KH2F1-RNAi. Full-amplicon analyses in this study, including the number of replicates, and *P*-values are also included (Supplementary Figures S3 and S4; Supplementary Table S2). (**I**) KREH2-RNAi enhanced MPSs. Dotted lines score the fold increase in NC/C ratio upon KREH2-RNAi at the site just 5′ of each intrinsic MPS (diamonds). When comparing two sets of samples (conditions), one set included at least three biological replicates and the other at least two, enabling *P*-value, average and SD calculations. ±Tet also labeled as ±T.

conceptually equivalent to (but determined differently from) exacerbated pause sites (EPSs) in similar studies by the Read lab (12). Either terminology indicates that misediting significantly increases just 5′ to the last correct edit (e.g. due to RNAi or in different stages). In summary, we showed that the observed KREH2-mediated changes in editing fidelity are enhanced on A6 in native RESC versus total mtRNA. We also provided the first evidence of KREH2-mediated differential effects on editing fidelity of A6 in PCF versus BSF cells. These differential effects target many positions along the A6 fragment examined, including the cluster of sites 31–38 across the initiator gRNA-1 in the 3′-UTR.

KREH2 differentially controls the formation of an abundant non-canonical sequence element in the A6 3′-UTR in PCF and BSF cells

Differences in relative A6 editing fidelity upon KREH2 knockdown between RESC and total mtRNA (37), and between PCF and BSF cells, potentially involve up- or down-regulation of canonical edits, non-canonical edits or both. In line with our prior study of KH2F1-RNAi (37), KREH2-RNAi in mtRNA significantly increased A6 non-canonical editing at most sites examined but did not seem to affect canonical editing (Supplementary Figure S3E, F). Notably, KREH2-RNAi in RESC significantly increased non-canonical editing largely in the 3′ terminus, including the

31–38 cluster. Conversely, canonical editing decreased in some 3′-terminal sites upon RNAi (Supplementary Figure S4E, F). We decided to examine all editing events in A6, particularly within sites 31–38 (marked by black arrowheads) in the 3′-UTR, which included high NC/C ratios in all our samples (Figure 4A). Notably, most sites in the 31–38 cluster included a specific non-canonical read at an exceptionally high frequency (>90% of all non-canonical reads on average; Figure 4B). These dominant non-canonical edits were >2-fold higher in BSF versus PCF mtRNA (32.5% versus 12.6% on average; $P < 0005$). The latter percentage values and subsequent ones are relative to all possible read types scored = 100%). In PCF cells, these non-canonical edits were enriched in RESC versus mtRNA (29.7% versus 12.6% on average; $P < 0.0005$) and further enriched by KREH2-RNAi in both RESC (to 36.7% on average; $P < 0.0005$) and mtRNA (to 17.1% on average; $P < 0.0005$). In contrast, in the BSF, KREH2-RNAi decreased these dominant non-canonical reads in mtRNA (to 29.3% on average; $P < 0.0005$).

We hypothesized that the above dominant non-canonical reads co-exist in the same molecules. The most frequent non-canonical reads at sites 31–38 predicted a consensus sequence element, with sites 35 and 36 having some variation in U insertion. Searches of amplicons that contain this consensus sequence while allowing any U number (n) at sites 35 and 36 revealed two top variants of this 3′-HFE in all samples (Figure 4C). Searches using either the 3′-HFE long version sites 31–38 or short version sites 31–34 produced the same total number of 3′-HFE-containing amplicons in each sample. This suggested that the short and long versions of the 3′-HFE are installed concurrently in A6.

We asked whether KREH2 differentially controls the frequency of the 3′-HFE in A6 in PCF and BSF cells and if changes in the 3′-HFE were more robust in RESC-bound transcripts. To this end, we determined the percentage of reads that contain the 3′-HFE in all samples (Figure 4D). In PCF cells, the 3′-HFE level in A6 was higher in RESC versus mtRNA (20% versus 7%, respectively, $P < 0.005$), and KREH2-RNAi significantly increased the 3′-HFE in both RESC and mtRNA. In this knockdown, the 3′-HFE level was about twice more in RESC versus mtRNA (36.1% versus 16.6%, respectively; $P < 0.0005$). An analysis using KH2F1-RNAi in PCF cells showed a similar phenotype to that observed with KREH2-RNAi. However, the 3′-HFE level in RESC-bound A6 was even higher in KH2F1-RNAi versus KREH2-RNAi (47.6% versus 36.1%, respectively, $P < 0.005$). This more robust phenotype in the KH2F1 knockdown was in line with the known KH2F1-dependent stabilization of KREH2 in PCF cells and concurrent loss of both proteins upon KH2F1-RNAi (37). In contrast to PCF cells, our analyses in BSF mtRNA showed the opposite phenotype. Namely, KREH2-RNAi significantly decreased the 3′-HFE level in A6 in this life cycle stage. These results confirmed that the generation of the 3′-HFE in A6 is enhanced by KREH2 knockdown in a PCF-specific manner, particularly in the context of the RESC complex.

We scored canonical reads in the first editing block guided by initiator gRNA-1, which showed a small decrease (not significant) in RESC-bound A6 from PCFs but not in other samples upon KREH2 knockdown (Figure 4E). We also compared the percentage of 3′-HFE vs. all other types of reads across this first block, including canonical, pre-edited and remaining non-canonical 'partial' (i.e. not matching the consensus 3′-HFE) (Figure 4F). We determined the percentage of 'partial' editing reads by subtracting from the total reads the sum of other read types in block 1: canonical, consensus 3′-HFE and pre-edited. In PCF cells, the increase in 3′-HFE reads upon knockdown of KREH2 was primarily linked to a loss of 'partial' reads in RESC and total mtRNA ($P < 0.05$ and $P < 0.005$, respectively). Pre-edited mRNA reads decreased slightly in RESC and mtRNA ($P < 0.25$ and $P < 0.1$, respectively). Block-1 canonical editing showed a small decrease in RESC ($P < 0.5$) but not mtRNA upon KREH2 knockdown. In BSF cells, a reduction in 3′-HFE reads upon KREH2 knockdown appeared linked to a moderate increase in pre-edited reads ($P < 0.03$) and a small decrease in canonically edited block-1 ($P < 0.2$) reads. The KREH2-RNAi-mediated loss of reads containing a 3′-HFE in BSF cells may reflect a general loss of total editing action in this stage. The above results support a model whereby generation of the non-canonical 3′-HFE in A6 involves active editing on RESC. Overall, KREH2-RNAi differentially affected the steady-state level of 3′-HFE-bearing A6, i.e. it increased in PCF but decreased in BSF cells. Our data provide the first example where loss of an editing protein, KREH2, up-regulates a specialized form of non-canonical editing in an mRNA substrate and does so specifically in the PCF stage. This KREH2-RNAi-mediated increase of the 3′-HFE primarily occurs in RESC-associated A6, so the effect is *in trans*.

The non-canonical 3'-HFE hinders A6 canonical editing, and its formation is directed by a novel putative regulatory gRNA

To better understand how the 3'-HFE is created and may affect editing in other A6 sites, we initially examined the top 10 most common amplicons derived from RESC-bound transcripts that carry the 3'-HFE long version ± KREH2-RNAi. Notably, in the highest frequency amplicon, all editing action had ceased precisely at the 5' end of the element sequence at site 38 (Figure 5A). Other amplicons contained junctions of non-canonical editing upstream of the 3'-HFE. Searches of the 3′-HFE long version found the same top amplicon species in all PCF and BSF samples ± KREH2-RNAi. All samples except for one had the same second top amplicon species with the last edit at site 39. (Supplementary Figures S5 and S6). A tally of the top 100 amplicons confirmed that the last edit occurs at the 5' end of the 3'-HFE or one site upstream in most samples (Figure 5A; each last edit site was tallied, and its percentage given above the sequence). Tallies of the 3'-HFE short version also showed the same top amplicon and position and percentage of the last edit, at site 38, in these molecules (Figure 5B). Other common amplicons in the 3'-HFE short version searches had their last edit at the end of the short element or nearby, including at site 39. Amplicons with the 3'-HFE short version typically contained the long version. These results indicated that the same molecular event creates the 3'-HFE short and long versions. The 3'-HFE would prevent anchoring by gRNA-2 and, thus, subsequent gRNAs in the

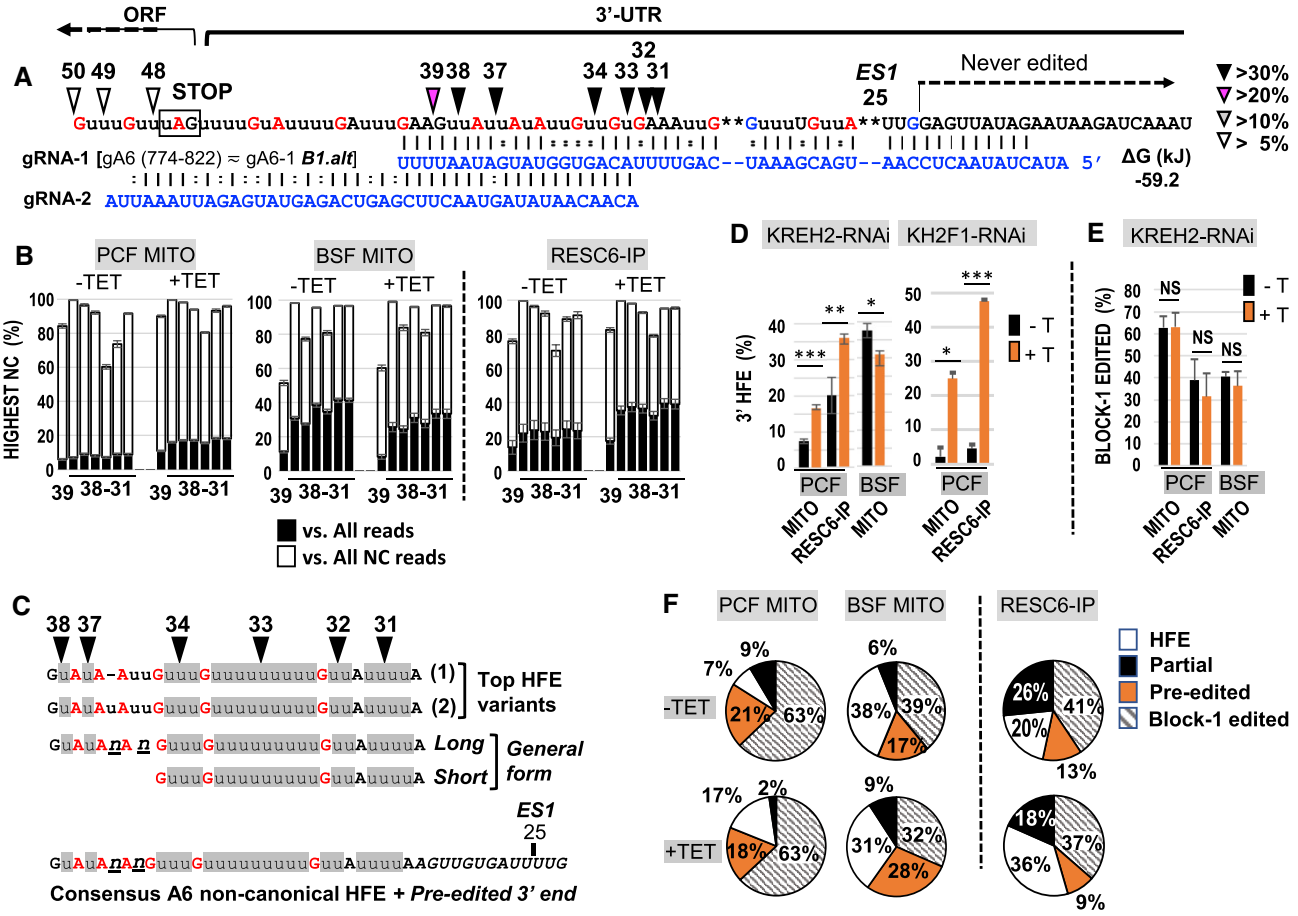


Figure 4. An abundant non-canonical 3′-HFE sequence in A6 in PCF and BSF cells, and its modulation by helicase complex REH2C proteins. (**A**) Canonically edited A6 3′ terminus. Color-coded letters are just 3′ of sites for sites requiring: insertion (red), deletion (blue) or changes (black). The ORF, 3′-UTR and never-edited regions are indicated. The first editing site (ES1) is at position 25 counting from the 3′ end. Illumina sequenced gRNA isoforms: gRNA-1 (gA6-1 *B1.alt*) in strain LISTER 427 (20) and gRNA-1 gA6 (774–822) in strain EATRO 164 (9) exhibit identical guiding function at block 1, and predicted gRNA-2 m0_306(II)_gA6_v2 [724–766] in strain EATRO 1125 (8) at block 2, produced the best match with canonically edited A6 (20,37). Color-coded arrowheads indicate sites that contain a dominant NC read representing >30% (black) of all reads in RESC + Tet (see below) or lower (see the color scale). (**B**) Actual percentage of the dominant NC read at each indicated site versus all reads (black) or versus all NC reads (white) in PCF mtRNA (Mito) or RESC, and BSF mtRNA ± KREH2-RNAi. The indicated range 38–31 includes only sites with dominant NC reads (>30% or >20%) in (A). (**C**) The 3′-HFE made by the dominant NC reads at sites 31–38. The top two 3′-HFE isoforms found in all PCF and BSF samples examined show the dominant NC reads in gray. General 3′-HFE long or short forms, where 'n' represents any T number at sites 35 and 36. Bottom: ∼42 nt extended 3' element, including the 3′-HFE and 3′-terminal pre-edited sequences. (**D**) Frequency of 3′-HFEs in PCF mtRNA or RESC, and BSF mtRNA in the indicated KREH2 or KH2F1 knockdowns. (**E**) Frequency of canonically edited block-1 in PCF mtRNA or RESC and BSF mtRNA ± KREH2-RNAi. (**F**) Frequency of 3′-HFEs, canonically edited block-1, pre-edited and other NC ('partial') reads in PCF and BSF ± KREH2-RNAi. When comparing two sets of samples (conditions), one set included at least three biological replicates and the other at least two, enabling *P*-value, average and SD calculations. ±Tet also labeled ±T.

canonical cascade. We have not found alternative gRNAs so far accounting for the most common 5′ junctions. Also, a search for canonical blocks 3 or 4 failed to find matches among all 3'-HFE-containing amplicons in our samples. These findings are consistent with the absence of suitable putative gRNAs in searches presuming progression from the 3′ element (long or short version) into upstream canonically edited sequence (Supplementary Figure S7).

Notably, 3'-HFE-bearing molecules carry a pre-edited 3' terminus (sites 25–30), including the first six positions in the A6 editing domain. In the canonical pattern, the first five positions require editing [sites 25–29; also known as ES1–ES5 (35,37)]. So, this extended element in A6 comprises a 3'-pre-edited terminus followed immediately by the non-canonical 3'-HFE. The consensus sequence of this abundant ≥42 nt element (sites 25–38) is: 5′-GuAuAnAnGuuu GuuuuuuuuGuuAuuuuAAGUUGUGAUUUUG-3'.

Multi-sequence alignments of the top 100 amplicons that carry the 3'-HFE confirmed the presence of the ≥ 42 nt extended 3' element with minor differences. This observation suggests that this extended element derives from a specific molecular event. We noted that the length of the extended 3' element suits the combined average sizes of the guiding and anchor regions in a typical gRNA (i.e. 20–40 nt and 6–11 nt long, respectively; (7,8). We hypothesized that one or more non-canonical gRNAs (i.e. gRNAs not matching the canonical pattern) might direct the formation of this 3' element in A6. A search in the essentially completely annotated minicircle genome from *T. brucei* strain EATRO 1125 (7,8) identified non-canonical gRNA

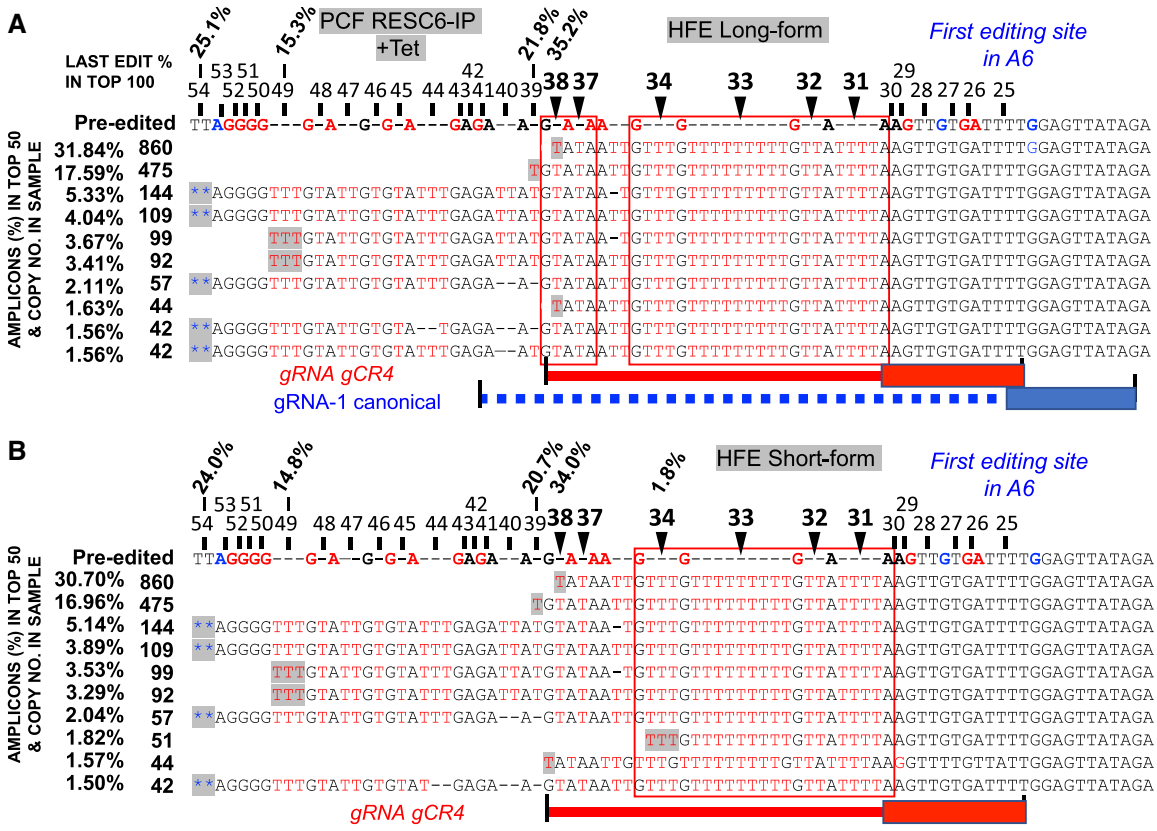
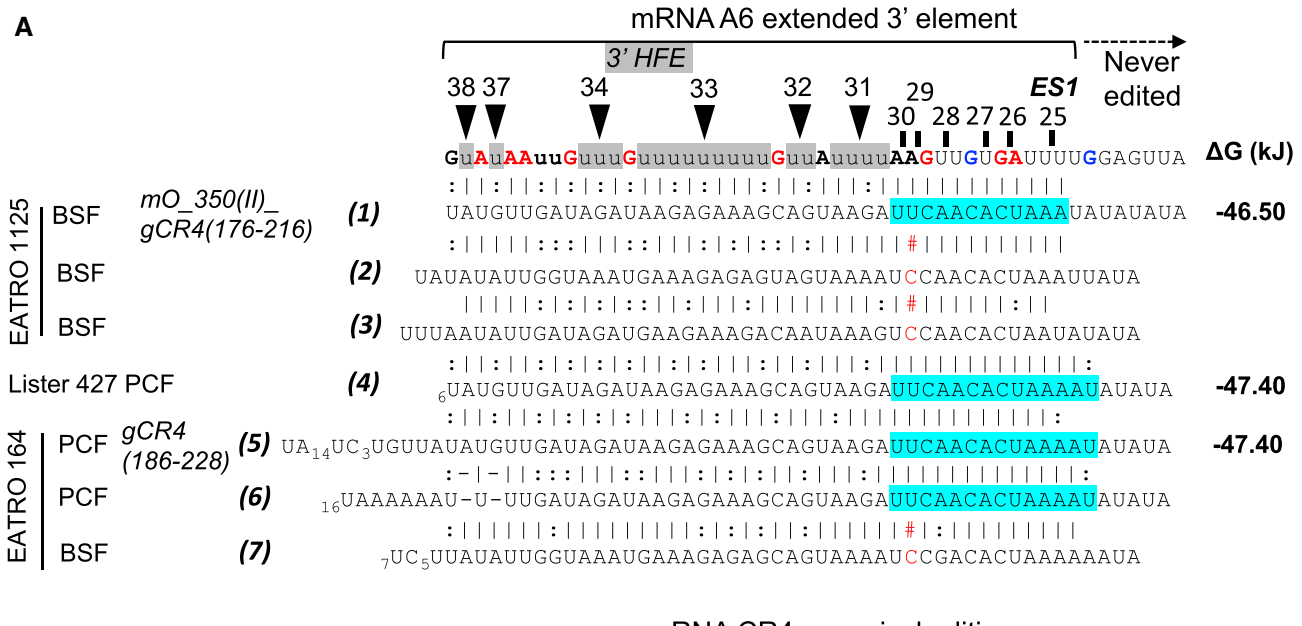


Figure 5. Most frequent A6 amplicons that contain the non-canonical 3′-HFE sequence. Sequence alignment of the top 10 amplicons in a representative sample of RESC6-IP KREH2-RNAi + Tet from PCF cells in searches of the (**A**) long or (**B**) short form 3′-HFE (sites 31–38 or sites 31–34, respectively). The last edit (gray) in each unique sequence is indicated as a percentage in the top 100 amplicons. Sequence 5′ to the last edit is completely pre-edited or includes a non-canonical editing junction of variable length. Sequence 3′ to the HFE is pre-edited in most amplicons. The top 10 amplicons in other samples were also examined (Supplementary Figures S5 and S6). The number of copies of each amplicon type in the representative sample and its percentage in the top 50 amplicons in that sample are shown. Dominant non-canonical edits in the 3′-HFE (boxed) are annotated as in Figure 4C. Searches of the short form produced amplicons with long form 3′-HFE. The anchor (box) and guiding region (dashed line) for canonical initiator gRNA-1 in blue and anti-initiator gRNA in red matching the 3′-HFE (straight line) are depicted. The first canonical editing site in A6 (position 25) is annotated in the sequence. Color-coded letters indicate sites requiring canonical insertion or deletion, as in Figures 2–4.

isoforms that match the extended 3’ element in A6 (Figure 6A; transcripts 1–3). One isoform matched the entire element except for one of four uridines at the first canonical site in pre-edited A6 (also known as ES1) (transcript 1). Surprisingly, this gRNA is the previously classified canonical gRNA mO_350(II)_gCR4(176–216) for CR4 mRNA editing progression (Figure 6B) (7,8). We also found equivalent isoforms in Illumina-sequenced RESC-bound and total mtRNA transcripts of the PCF strain Lister 427 used here (transcript 4) (20), and in total mtRNA of strain EATRO 164, including gCR4(186–228) (transcripts 5–7) (10). In strain Lister 427, the most common RESC-bound gRNA isoform (transcript 4) and one isoform in EATRO 164 (transcript 5) both have a 13 nt anchor that completely matches the pre-edited 3′ terminus via Watson–Crick base pairing, including all four uridines in the first canonical site in pre-edited form (also known as ES1) in A6. A 30 nt ‘guiding’ region in most isoforms (Figure 6A) precisely matches the non-canonical 3’-HFE (including G–U wobble base pairing). Similar isoforms were additionally found in PCF strains TREU 667 and TREU 927 (Donna Koslowsky, personal communication). The C/T polymorphism and mismatch in some identified isoforms could cause alternative non-canonical insertion at site 30 (i.e. +4U). However, site 30 exhibited low NC/C values (Figure 3), suggesting that these isoforms are infrequently or not utilized.

The conservation of gRNA isoforms above in multiple *T. brucei* strains and their full complementarity to the extended 3’ element suggest that these gRNAs are biologically relevant to A6 editing. Notably, available Illumina data in total mtRNA from PCF and BSF strains Lister 427, EATRO 164 and strain EATRO 1125, and in RESC from PCF strain Lister 427, consistently indicated a lower copy number of the isoforms examined for canonical A6 initiator gRNA-1 versus gRNA mO_350(II)_gCR4(176–216) (Figure 6C; and data not shown) (8–10,20). The above differences in gRNA copy number suggested that KREH2 enables preferential utilization of the rare canonical gRNA-1 over the more abundant non-canonical gCR4 in the A6 target. However, KREH2 disruption removes the constraint on gCR4, diminishes canonical gRNA-1 function or both. Either way, KREH2 knockdown in the PCF stage



C

	Number of gRNA transcripts in library			
	LISTER 427	EATRO 164		EATRO 1125
	PCF	PCF	BSF	BSF
Initiator gA6 isoforms	1	1	ND	6,501
mO_350(II)_gCR4(176-216) isoforms	320	7	166	65,332

Figure 6. A novel putative regulatory gRNA may install the extended 3′ element in A6. (**A**) Sequence alignment of the mRNA A6 extended 3′ element, which includes the 3′ HFE (annotated as in Figure 4C) with strain EATRO 1125 gRNA mO_350(II)_gCR4(176-216) (transcript 1) and isoforms (transcripts 2-3) in strain EATRO 1125 (8), and isoforms in strains Lister 427 (used in this study; transcript 4) (20) and EATRO 164 (transcripts 5-7) (9). The anchor region (blue) in all isoforms matches the first six sites in A6 in pre-edited form. Isoforms in strain Lister 427 and EATRO 164 (transcripts 4 and 5-6, respectively) match all four Us in the first editing site (aka *ES1*, position 25). Some anchors have a mismatch (red) or wobble base pair. One isoform specifically guides for short-form 3′-HFE (transcript 6). (**B**) Sequence alignment of gRNA mO_350(II)_gCR4(176–216) with its cognate mRNA CR4, including the anchor region highlighted (blue). These findings imply a novel gRNA type that installs the abundant 3′ element in A6. This gRNA may serve a dual editing role: non-canonical in A6 and canonical in CR4. The gRNA transcripts shown were Illumina sequenced (8,9,20). A number at the 3′ terminus indicates post-transcriptionally added Us. (**C**) Number of gRNA transcripts in available databases in PCF and BSF stages. Canonical initiator gRNA-1 isoforms in EATRO 164, gA6 (774–822) and Lister 427, gA6 *B1.alt*, have an identical guiding capacity to the initiator in EATRO 1125. Copy number of EATRO 164 and EATRO 1125 gRNAs was determined in total mtRNA; copy number of Lister 427 gRNAs was determined in RESC6-IPs (8,9,20; and available alignments online). Not detected (ND).

specifically increased non-canonical gCR4 targeting of A6. Overall, the Lister 427 version of mO_350(II)_gCR4(176–216) may represent a KREH2-modulated bifunctional gRNA, i.e. capable of opposing dual roles in either canonical CR4 editing or non-canonical A6 editing. KREH2 differentially modulates this gCR4 targeting on A6 in PCF and BSF stages. Remarkably, gCR4 anchor and continuous duplexes with the 3′ element-containing A6 exhibit lower ΔG values than with CR4 mRNA (Figure 6A, B), suggesting that non-cognate gCR4–A6 mRNA pairs are thermodynamically more stable than cognate gCR4–CR4 mRNA pairs. This argues against random off-targeting of mO_350(II)_gCR4(176–216) isoforms on A6 but rather suggests a specific and fixed event. We confirmed that induction of KREH2-RNAi knockdown inhibits full editing of mRNA CR4 (Figure 6C), as with other mRNAs examined in this study (Figure 1C). Our assay detected full editing of CR4 in PCFs but not in BSF cells.

Further supporting a proposed bifunctional role of gCR4 gRNA, we readily isolated *in vivo* generated chimeric molecules of gCR4 with both A6 and CR4 mRNAs. Chimeras may not be true editing intermediates but validate gRNA–mRNA pairing *in vivo* (13,14,53,54). Remarkably, both chimeras appeared to use the same gCR4 isoform (4) in strain Lister 427, based on their sequenced 3′ terminus (Figure 7). Thus, KREH2 is the first example of a specific editing factor that controls a non-canonical gRNA, which appears to be bifunctional.

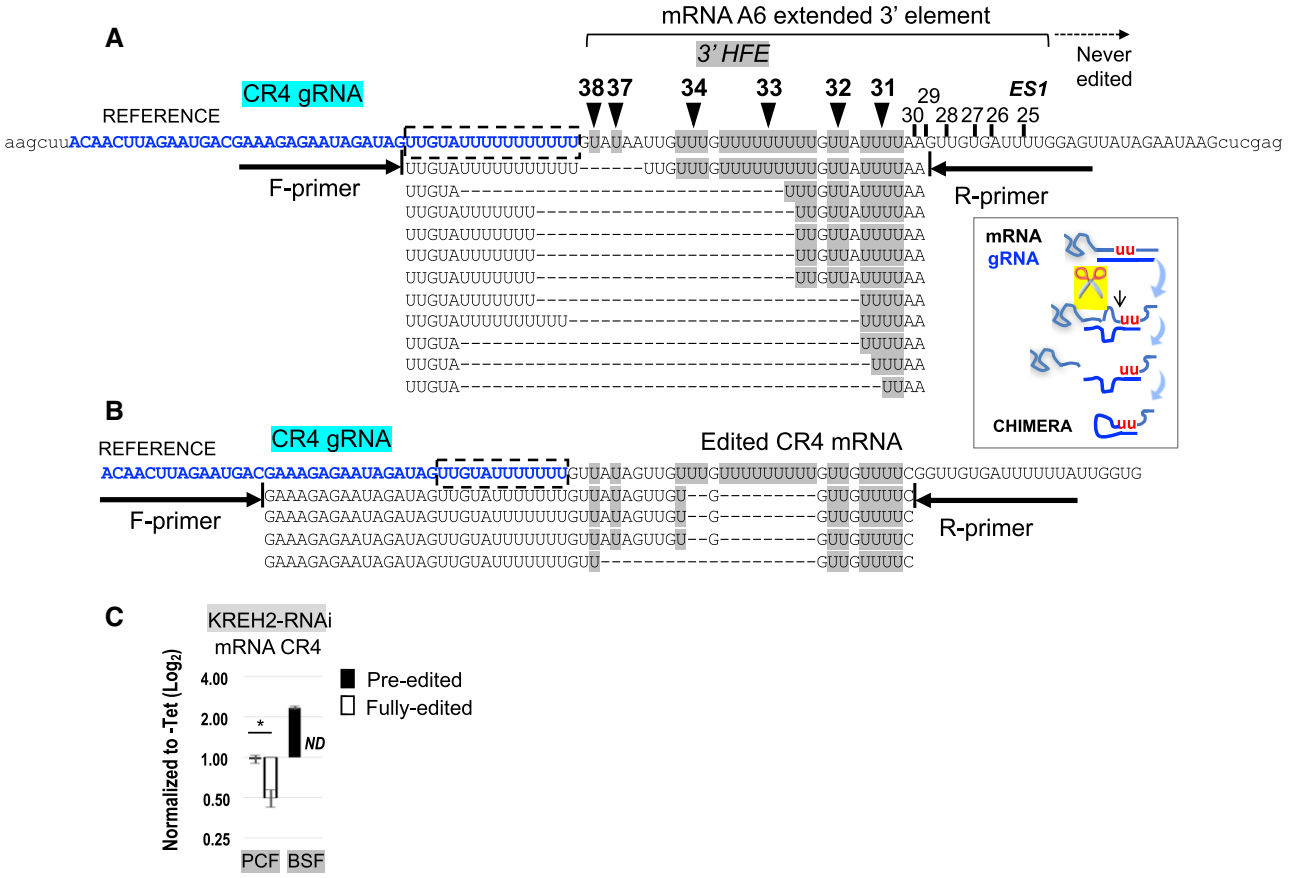


Figure 7. *In vivo* chimera formation of the putative bifunctional gRNA gCR4 with mRNAs A6 and CR4. Multisequence alignment of RT-PCR-amplified chimeras between gRNA gCR4 and (**A**) HFE-bearing A6 or (**B**) canonically edited CR4 mRNA. The top sequence is a reference of the predicted chimeras using the gCR4 3′ terminus of isoform 4 (Lister 427 in Figure 6). gRNA in blue with identical 3′-terminal bases plus U-tail captured by Sanger sequencing in both chimeras (dotted box). Common U-insertions in HFE and CR4 mRNAs (gray shade). Forward (F) and reverse (R) primers. A drawing of *in vivo* chimera formation depicts mRNA cleavage and subsequent ligation of the newly excised mRNA 5′ end with the 3′ terminus of the hybridized gRNA. (**C**) RT-qPCR of mRNA CR4 pre-edited or fully edited in PCF and BSF stages upon KREH2-RNAi. Fully edited mRNA was not detected (ND) in the BSF.

KREH2-dependent editing control by a 3′ element-associated 'repressive' RNA structure

A6 targeting by gRNA mO_350(II)_gCR4(176–216) covers most sites typically modified by canonical editing initiation in the A6 3'-UTR. However, ablation of editing at the first few sites of A6, just downstream of the 3′-HFE, may not be explained by the gCR4 gRNA anchor hybridization alone. The anchor region for canonical initiator gRNA-1 remains intact in the never-edited region of A6 (Figure 4A), so this gRNA could potentially direct editing of the first few sites in A6. Even after the extended 3′ element has been installed, A6 could be potentially 'repaired' by canonical initiator gRNA-1 in later rounds of editing. To explain this conundrum, we reasoned that changes in RNA secondary structure might also disrupt initiator gRNA-1 function. We addressed this possibility by experimentally determining the secondary structure of A6 via DMS mutational profiling with sequencing (DMS-MaPseq) (Figure 8A-C). This RNA structure probing strategy takes advantage of a high-fidelity processive thermostable group II intron reverse transcriptase (TGIRT) enzyme (45). DMS rapidly and specifically labels the Watson–Crick face of open and accessible adenine and cytosine bases in the RNA. This probing strategy is suitable for editing mRNAs because their sequence is purine rich (Supplementary Figure S1) (8,9).

We examined full-length *in vitro* T7-transcribed A6, either pre-edited or bearing the most frequent extended 3′ element in all samples (Figure 5; Supplementary Figures S5 and S6). Notably, DMS reactivity profiles showed an evident decrease in signal across sites 31–38 in the 3′-HFE sequence (Figure 8C). A particularly low DMS signal in the 3′-HFE sequence suggests that this region has lost flexibility and forms a highly stable duplex. On the other hand, a similar DMS reactivity in the remaining A6 sequence—pre-edited mRNA or bearing the non-canonically edited extended 3′ element—suggested that sequences outside the 3′-HFE are less affected. We used the reactivities as folding constraints to generate structural models of A6 (Figure 8A, B). These models showed that the ∼42 nt extended 3′ element forms a highly stable structural domain. The overall structure of this 3′ element-containing A6 isoform is ∼32% more stable than pre-mRNA (predicted ΔG –42.2 kJ and –32.1 kJ, respectively). Our experimentally determined structures support a model whereby the loss of initiator gRNA-1 function in 3′ element-bearing A6 molecules

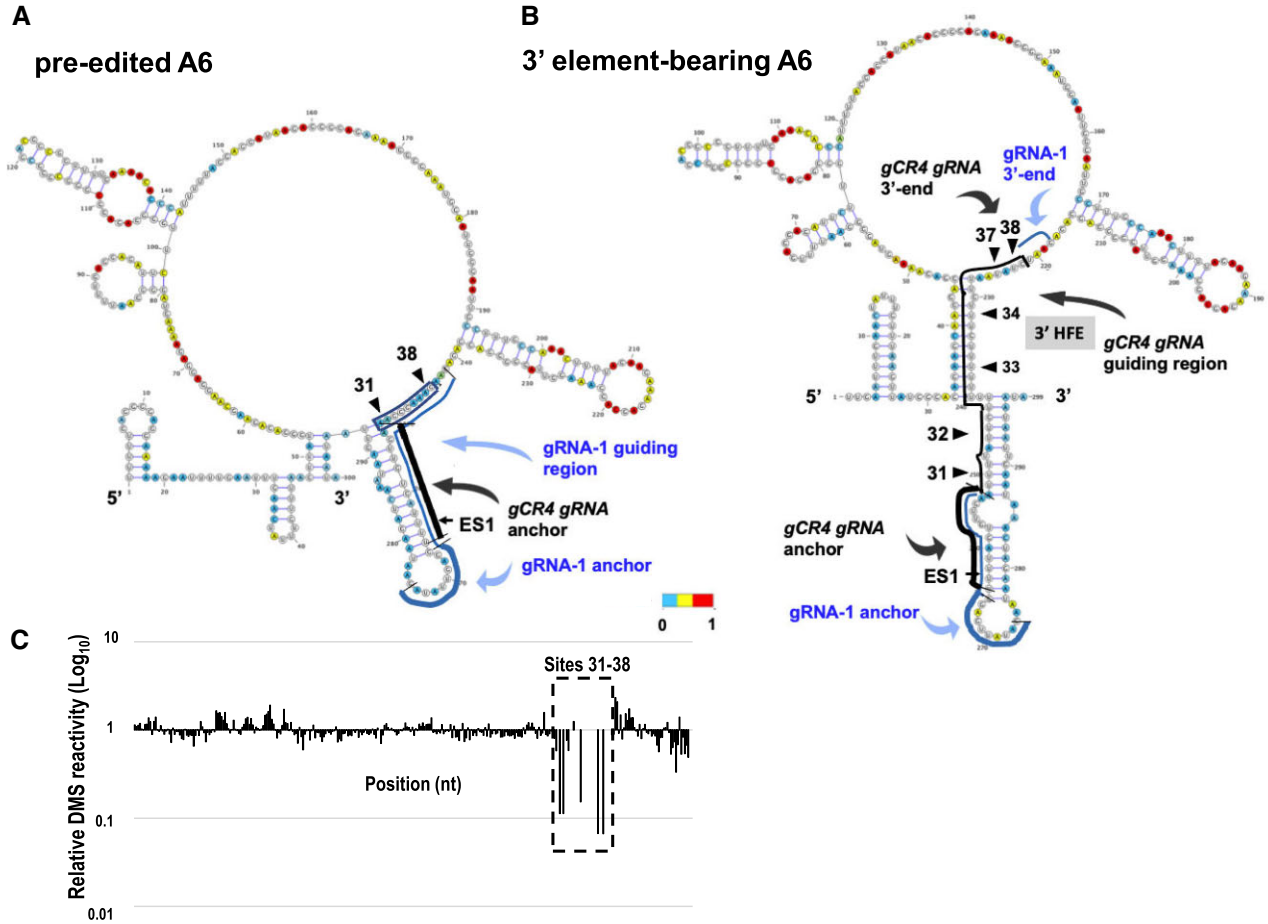


Figure 8. Experimentally determined structure of 3′-HFE-containing full-length A6 and analyses of mRNA RESC association. DMS-MaPseq secondary structure of *in vitro* transcribed full-length A6: (**A**) pre-edited (403 nt) and (**B**) one of the top two 3′-HFE-containing A6 isoforms (426 nt) in all samples. Diagrams depict 300 nt on the 3′ end of each construct. Sites 31–38 with non-canonical edits in the 3′-HFE and/or sites matching guiding and anchor regions of anti-initiator gRNA (black line) and canonical initiator gRNA-1 (blue line) are annotated. Nucleotides are colored by DMS reactivity calculated as the average ratiometric signal per position across two biological replicates normalized to the highest number of reads in the displayed region, which is set to 1.0. (**C**) Relative DMS reactivity at each nucleotide normalized to the signal in the pre-edited molecule for the same nucleotide. The dotted black box indicates sites 31–38 which had the greatest loss in DMS reactivity for any region in the 3′-HFE-bearing molecule.

is due to a 'repressive' RNA structure involving the element itself. Overall, KREH2-modulated non-canonical action by gCR4 gRNA might prevent canonical gRNA-1-mediated repair of 3′-HFE-containing transcripts, by forming a repressive RNA fold that sequesters the 3'-UTR in A6.

DISCUSSION

Most mRNA molecules in mitochondria carry non-canonical events in a junction region, whilst a minor number complete canonical editing. An open question is whether non-canonical sequences get fixed, and at least some may serve specific functions (12,37,42). Here, we report the first identified editing protein, RNA helicase KREH2, that differentially modulates non-canonical editing. KREH2-RNAi knockdown affects the general A6 pool in at least two ways: (i) it enhances natural pausing, including at major pause sites, during 3′–5′ editing progression involving canonical initiation; and (ii) it enhances abundant programmed (i.e. gRNA-directed and regulated) alternative editing in the 3′-UTR, without involving canonical initiation, that we characterized in more detail. We showed that KREH2 down-regulation does not evidently affect the stability of other editing proteins. However, we cannot rule out KREH2 affecting RESC1/2 interactions since this RNA helicase controls editing fidelity in RESC-bound mRNA and associates with RESC complexes via RNA (20,21,34,35,37). We are currently examining this possibility.

A6 maturation occurs in both PCF and BSF stages; however, KREH2 depletion differentially controls installation of a 3′-HFE in A6. This 3′-HFE exhibits at least three prominent features: first, an exact sequence match with a proposed regulatory gCR4 gRNA identified in this study; second, PCF-specific up-regulation of formation of the 3′-HFE induced by KREH2-RNAi; and third, the establishment of a repressive RNA structure by the 3′-HFE that may sequester the 3'-UTR, hindering canonical editing of HFE-containing A6. This repressive RNA fold may occlude access to canonical A6 initiator gRNA-1, preventing 3′ element removal via potential proofreading editing. Traditional 3′–5 progression involving canonical initiation

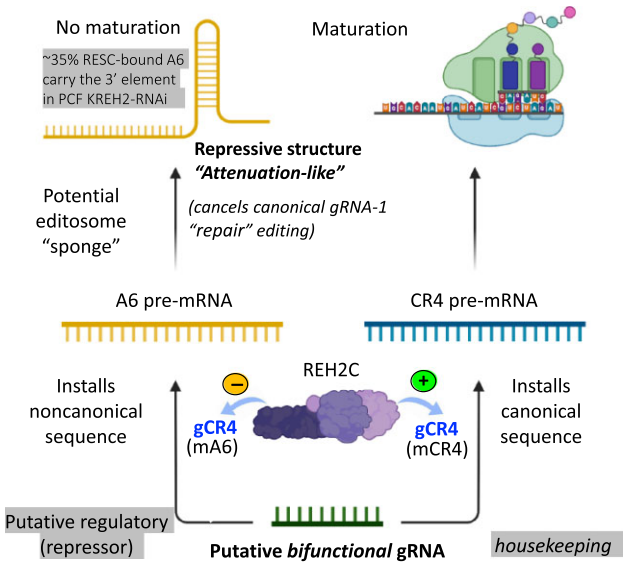


Figure 9. Model of REH2C-dependent developmental control of non-canonical editing by a putative bifunctional gRNA in A6. Amidst extensive RNA editing of unclear function, KREH2, a REH2C-associated helicase, controls programmed (i.e. gRNA-directed and regulated) non-canonical editing. A novel putative bifunctional gCR4 gRNA directs non-canonical editing to generate a 3′ element in A6 besides serving canonically in CR4 editing. This novel gRNA-directed function on A6 seems most active in RESC and is modulated differently by KREH2 during the life cycle. In the PCF, REH2C negatively controls gCR4 gRNA action on non-cognate A6. Loss-of-function mutants (KREH2 or KH2F1 knockdowns) up-regulate the 3′ element to an astonishing ∼35% of RESC-bound A6 molecules. REH2C may positively control gCR4 gRNA action on its cognate target, as is generally expected for canonical editing. The 3′ element forms a stable structure that sequesters the 3′-UTR and blocks canonical editing initiation-mediated 'repair' of the 3′-UTR, representing a new type of attenuation. In BSF cells, this element is particularly abundant (∼40%) but not up-regulated by KREH2-RNAi. Instead, KREH2 may be required to maintain the 3′ element at high levels in BSF cells. However, A6 maturation occurs in both stages, but apparently at different levels. Thus, KREH2 differentially controls a novel potentially regulatory gRNA, a putative repressor, preventing excessive A6 maturation in mitochondria. Overall, the current study identified the first example of programmed non-canonical editing, which is both mitochondrial genome encoded and regulated and may control mitochondrial physiology during development. Our model suggests that at least some mitochondrial non-canonical editing became fixed and regulated in the long evolutionary history of kinetoplastids.

(i.e. without the 3′-HFE) would explain the observed KREH2-RNAi effects on editing pausing in a fraction of the A6 pool. Key observations in the current study leading to this proposed model of KREH2-RNAi-mediated control of A6 editing (Figure 9), with differences in PCF and BSF stages, are discussed next.

KREH2 negatively controls a proposed regulatory gRNA in PCF A6 mRNA

The 3'-UTR of A6, covered by the first two gRNAs, exhibits particularly high total editing, and we have hypothesized that KREH2 might regulate editing in this region (37). Notably, >60% of RESC-bound A6 molecules in PCF cells contained either non-canonical editing of the 3′ element directed by gCR4 gRNA or canonical editing of block-1 by initiator gRNA-1 (20% versus 41%, respectively). So, these two gRNAs combined account for most total edits across the A6 3' terminus. The 3′ element formed by gCR4 gRNA was enriched in RESC versus total mtRNA, just as we reported for canonical editing of several substrates (20,37). Also, in RESC-bound A6, KREH2-RNAi induction augmented the 3′ element to a frequency comparable with that of canonically edited block-1 (36% versus 37%, respectively). Combined, these two sequences accounted for >70% of all edits at the first few sites in RESC-bound A6 upon KREH2 knockdown. Thus, KREH2-RNAi in PCF cells significantly up-regulated gCR4 gRNA function but slightly down-regulated canonical initiator function (at least in RESC-bound A6). Together, the extensive gCR4 gRNA action at the A6 3'-UTR, its up-regulation by KREH2 knockdown and the conservation of gRNA isoforms in five strains of *T. brucei* examined indicate that this proposed novel gRNA type is biologically relevant. By installing RNA structure, gCR4 gRNA-directed editing potentially prevents canonical gRNA initiator removal of the 3′ element. In this particular situation, gCR4 gRNA could act as an anti-initiator.

Even if KREH2 knockdown only moderately affects canonical initiation, it inflicts a cumulative inhibitory effect over the entire cascade, as shown by RT-qPCR or RNA-seq. The canonical gRNA-1 and gCR4 gRNAs may not necessarily be exclusive. However, it seems that the latter may serve to keep much of the A6 substrate in a repressed state. Alternatively, canonical initiator and gCR4 gRNAs might compete for A6 substrate, but scored levels of edited A6 at steady state could reflect differences in stability or modification of transcripts after canonical or non-canonical editing. Similarly, the steady-state level in edited mRNA often changes substantially without corresponding changes in pre-edited mRNA. Differences in the stability of different species could also involve changes in AU 3′ tails (55), but additional studies are needed to address this possibility.

Selective utilization of a canonical initiator over non-canonical gRNAs in PCF and BSF cells

The relative abundances of canonical initiator and gCR4 gRNAs differ in the available PCF and BSF gRNA transcriptome analyses (Figure 6C) (20,56). The proposed regulatory gRNA mO_350(II)_gCR4(176–216) in A6 occurs at a generally higher frequency than canonical initiator gRNA-1 in all available libraries of PCF and BSF strains. The difference in cellular abundances of gCR4 and initiator gRNAs suggests that KREH2 restricts non-canonical utilization of gCR4 gRNA on A6. KREH2-RNAi knockdown in PCF cells would remove this restriction, causing an increase in gCR4 targeting of A6. As noted above, in RESC-bound A6, canonical initiator gRNA-1 action surpasses gCR4 gRNA action ∼2:1 in PCF cells (41% versus 20%, respectively). The reported copy number and our scored editing levels in RESC-bound A6 suggest a >300-fold higher utilization of canonical initiator gRNA than gCR4 gRNA in PCF Lister 427. This differential implies a built-in mechanism in the editosome holoenzyme, which specifically enhances A6 canonical editing initiation. The predicted anchor of pre-edited A6 substrate is more

stable with gCR4 than with canonical initiator gRNA (14 nt, 5′-UAAAAUCACAACUU-3' Tm = 32.4°C, and 12 nt, 5′-CUAUAACUCCAA-3', Tm = 27.3°C, respectively), further suggesting that specificity factors for canonical initiation are necessary. However, the duplex between canonical initiator gRNA-1 and canonically edited A6 product (Figure 4A) is more stable than that between gCR4 and HFE-containing A6. Such proposed initiation factors may increase the affinity of the initiator gRNA anchor 'seed' region and target mRNA in RESC. This concept is reminiscent of enhanced thermodynamic stability of the miRNA seed region and target mRNA in the Ago2-RISC microprocessor (57). The helicase activity in the REH2C complex may normally shift the free energy landscape restricting A6 binding to gCR4 gRNA. Such restriction is partially removed by disruption of helicase activity. An RESC protein subunit, RESC14 (also known as MRB7260), has been implicated in selective gRNA utilization in PCF cells (24,44) and may be a possible gRNA specificity factor.

Putative bifunctional and other possible types of alternative gRNAs

Alternative gRNAs have been proposed that may alter the coding potential of the edited transcriptome in *T. brucei* (7–9,20,42,43). This includes the generation of possible dual-coding genes, where alternative editing at mRNA 5' ends can alter the choice of start codon and the ORF (43). Also, many non-canonical gRNAs of unknown function were identified (7,8). A few non-canonical gRNAs may act as terminators in mRNA CO3 (44) by anchoring to canonically edited sequences and inserting a sequence that derails canonical progression. In contrast, the gCR4 gRNA identified in the current study anchors to pre-edited sequence installing a structural 3′-HFE that may help repress canonical editing in HFE-bearing A6. *In vivo,* gRNA–mRNA chimeras indicate gCR4 gRNA hybridization with HFE-containing A6 and cognate CR4 mRNAs. Our *in vivo* data further support a bifunctional nature of gRNA mO_350(II)_gCR4(176–216). In the kinetoplastid *L. pyrrhocoris*, some gRNAs were proposed to direct both canonical and non-canonical editing, including gRNAs in ND9 that may introduce non-canonical editing in RPS12 (40,58). However, these non-canonical pairs were identified under relaxed conditions not requiring strict anchor region Watson–Crick base pairing, and the anchor and/or guiding regions have multiple mismatches. Those gRNAs might alter the ORF without derailing upstream canonical editing.

We propose that the gCR4 gRNA utilization in A6 reported here is biologically important in *T. brucei* and other kinetoplastids for these reasons: a continuous duplex between the gCR4 gRNA and 3′-HFE in strain Lister 427, and the 3′-HFE extreme cellular abundance and its specific control by KREH2 in *T. brucei*. Also, gCR4 gRNA forms a thermodynamically more stable anchor with 3′-HFE-bearing A6 than its cognate CR4 mRNA, further supporting a bifunctional role. Thus, the gCR4 gRNA identified here is a proposed novel type of regulatory transcript in editing.

Differential gRNA utilization and hypothetical energy efficiency control by an abundant repressive element

As noted above, KREH2-RNAi knockdown in BSF did not up-regulate gCR4 gRNA utilization in A6 as in PCF cells. Instead, gCR4 gRNA action on A6 significantly decreased upon KREH2-RNAi knockdown in BSF cells. This developmental difference indicates that the helicase KREH2 restricts the use of this gRNA specifically in PCF cells. Notably, gCR4 gRNA function on A6 in total mtRNA is significantly higher in BSF than in PCF cells (38% versus 7%, respectively) in the absence of KREH2 knockdown. Accordingly, cumulative total editing and canonical editing were significantly lower in BSF versus PCF stages due to the 3' element hindering upstream editing ($P < 0.01$ and $P < 0.0005$, respectively). Because KREH2-RNAi did not reduce the level of gCR4 gRNA action on A6 in PCF cells, other editing factors in this stage may be needed to establish the abundant repressive 3' element in steady-state A6. The exceedingly high action by gCR4 gRNA on A6, particularly in BSF cells, may require high-affinity anchoring and a high copy number of this gRNA. However, these gRNA features do not necessarily correlate with the editing level of the mRNA target (7,35,56), so RESC components may increase the 'seeding' potential of gRNA gCR4 on A6, particularly in BSF cells. A high level of A6 'dead-end' molecules (i.e. bearing the repressive 3' element) in PCF and BSF stages, and its PCF-specific up-regulation by KREH2, suggests that trypanosomes purposely regulate non-canonical editing on a large scale (as discussed below). Editosomes may be recruited to introduce repressive editing in at least six sites (17 Us inserted) in the A6 3′ terminus, thereby saving resources otherwise required for canonical editing at all 185 sites (447 Us inserted and 28 Us deleted). The energy used in creating 'repressive' editing at the A6 3' terminus may be far less than the energy required to create mature transcripts. Early *in vitro* mechanistic studies defined that a full round of editing at each site entails three basic protein-catalyzed steps: mRNA endonuclease cleavage, U addition/removal and ATP-dependent ligation (53,54,59,60). Based on this basic reaction alone, maturation of A6 would consume ∼30 times more ATP than just installing the 3′-HFE. A general model invoking programmed non-canonical editing to control canonical editing and, thus, energy efficiency would be a novel feature of trypanosomal biology. Massive programmed non-canonical editing at the A6 3' terminus upon KREH2 knockdown could titrate factors needed for general canonical editing. Such putative editosome hijacking, or 'sponge effect', by A6 might contribute to the global editing phenotype observed upon RNAi. In that case, BSF and PCF stages could purposely modulate non-canonical 'repressive' editing in A6 to limit overall editing action in A6 and potentially all canonical editing during development.

Other mitochondrial RNA helicases include the DEAH-Box family member, KREH1, which participates in canonical editing (61,62), and a DExH-Box KREH2-paralog (Tb927.4.3020). Both helicase proteins were found in purifications of KREH2, but the interaction was RNase sensitive (34). KREH2 and its much shorter paralog share an identical domain organization (63,64). However, initial RNAi-induced knockdown studies of the latter

protein could not link it with canonical editing (Madina *et al*., personal communication; 35,63). It is conceivable that KREH1 or the KREH2-paralog may control programmed non-canonical editing during development, but additional studies are needed to examine this possibility.

Implications of programmed non-canonical editing on the constructive neutral evolution hypothesis and mitochondrial physiology

As an alternative to selectionist or adaptation views, RNA editing could have appeared by constructive neutral evolution (CNE) as it appears gratuitously complex, comprising more features than its basic function may demand (65–67). Under this hypothesis, non-canonical editing junctions of unpredictable composition and unclear origin occur without evolutionary benefits and are neutrally fixed without positive selection (i.e. by genetic drift). Our conclusion that, amidst extensive non-canonical editing, specialized events are, in fact, encoded and regulated by specific factors challenges the CNE view. However, one should distinguish between the initial evolution and fixation, which could have been neutral, and subsequent elaboration, including adapting new functions that could be beneficial. So, editing could have evolved via CNE, but it then represented a new playground for Darwinian-type evolution to take place.

A6 editing is presumably essential in both BSF and PCF stages; however, different levels may be required during development. Indeed, BSF cells exhibited less total edited A6 than PCF cells in the fragment examined (Figure 2). This developmental difference in edited A6 was also detected but not discussed in a prior study (44). Differential A6 function levels and regulation of the repressive 3′-HFE may be expected due to the environmental changes and bioenergetics adaption during development. PCF cells utilize a branched mitochondrion that is fully developed containing many cristae. The mitochondrion in BSF long slender cells is smaller and devoid of cristae. The F_OF_1-ATP synthase complex generates ATP in PCF cells (forward mode) but becomes a perpetual consumer of ATP in BSF cells (reverse mode) (68). Interestingly, a deficiency of F_OF_1–ATP synthase in PCF cells decreased ATP levels and cell growth; however, BSF cells withstand a substantial loss of the complex without evidently affecting cell growth (69). Finally, mutant dyskinetoplastic trypanosomes that lack kDNA are possible in BSF but presumably not in PCF cells (70,71), altogether underscoring a strict requirement for mitochondrial genome function and regulation in PCF but not in BSF stages.

In line with a reduced level of edited A6 in BSF versus PCF cells, a more considerable amount of repressive 3′-HFE may be needed to prevent wasteful energy use for editing in BSF cells where less mRNA maturation is needed. In this scenario, negative regulation of the 3′-HFE may not be required in BSF cells. In contrast, KREH2 probably acquired an additional role in repressing and modulating the formation of this element in PCF cells. If the 3′-HFE helps control global editing, as discussed above, our proposed scenario may suit the physiological needs of PCF and BSF stages. Overall, our findings may reveal evolutionary benefits of at least some non-canonical editing, fixed by positive selection and differentially regulated by editosomes. Finally, A6 regulation could also help modulate cellular ATP and ADP levels during cleavage and ligation in the full round editing reaction defined *in vitro* (54,72,73). The above possible regulatory effects of the 3′-HFE are speculatory. However, our finding of the 3′ element genesis and its developmental regulation in > 30% of A6 transcripts indicates natural selection, not stochasticity, of a specialized gRNA-directed event in the long evolutionary history of *T. brucei*.

Overall, we have identified the first editing protein, RNA helicase KREH2, that controls abundant non-canonical editing during development, which modulates pausing during traditional 3′–5′ editing progression, and a remarkably abundant 3′-HFE directed by a novel putative regulatory gRNA. Such a 3′-HFE forms a proposed repressive RNA fold that sequesters the 3’-UTR and prevents initiator gRNA action, which could remove the non-canonical 3′-HFE. *In vivo* chimeras of gCR4 gRNA with 3′-HFE-containing A6 or cognate CR4 mRNAs support a bifunctional role for this gRNA in *T. brucei* mitochondria. Our findings support a general model in which at least some non-canonical editing is fixed and part of novel molecular regulatory mechanisms in editosomes (Figure 9). Abundant programmed non-canonical editing by REH2C-controlled specialized gRNAs, including proposed gRNA repressors, may modulate ‘edited’ protein biogenesis and overall mitochondrial physiology during development.

DATA AVAILABILITY

RNA sequencing data are deposited in the NCBI SRA under BioProject ID PRJNA903173.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank members at the Institute for Genome Sciences and Society at Texas A&M University for their excellent assistance with deep sequencing. The Afasizhev lab kindly provided the antiserum against RESC1/2 (GAP2/1). The Read lab kindly provided the antisera against RESC13 (KRGG2). We are very grateful to Thavy Khem, Karissa Beauchemin and Nathan Parulian for their assistance in preparing Excel files and charts.

FUNDING

The National Science Foundation [1616845 to J.C.R. and 2140153 to S.M.]; TAMU X-grant [to X.Z., L.Z. and J.C.R.]. Funding for open access charge: National Science Foundation [1616845].
Conflict of interest statement. None declared.

REFERENCES

1. Kostygov,A.Y., Karnkowska,A., Votypka,J., Tashyreva,D., Maciszewski,K., Yurchenko,V. and Lukes,J. (2021) Euglenozoa: taxonomy, diversity and ecology, symbioses and viruses. *Open Biol.*, **11**, 200407.

2. Stuart,K., Brun,R., Croft,S., Fairlamb,A., Gurtler,R.E., McKerrow,J., Reed,S. and Tarleton,R. (2008) Kinetoplastids: related protozoan pathogens, different diseases. *J. Clin. Invest.*, **118**, 1301–1310.
3. Zimmer,S.L. (2019) Revisiting trypanosome mitochondrial genome mysteries: broader and deeper. *Trends Parasitol.*, **35**, 102–104.
4. Jensen,R.E. and Englund,P.T. (2012) Network news: the replication of kinetoplast DNA. *Annu. Rev. Microbiol.*, **66**, 473–491.
5. Lukes,J., Guilbride,D.L., Votypka,J., Zikova,A., Benne,R. and Englund,P.T. (2002) Kinetoplast DNA network: evolution of an improbable structure. *Eukaryot. Cell*, **1**, 495–502.
6. Stuart,K., Allen,T.E., Heidmann,S. and Seiwert,S.D. (1997) RNA editing in kinetoplastid protozoa. *Microbiol. Mol. Biol. Rev.*, **61**, 105–120.
7. Cooper,S., Wadsworth,E.S., Schnaufer,A. and Savill,N.J. (2022) Organization of minicircle cassettes and guide RNA genes in *Trypanosoma brucei*. *RNA*, **28**, 972–992.
8. Cooper,S., Wadsworth,E.S., Ochsenreiter,T., Ivens,A., Savill,N.J. and Schnaufer,A. (2019) Assembly and annotation of the mitochondrial minicircle genome of a differentiation-competent strain of *Trypanosoma brucei*. *Nucleic Acids Res.*, **47**, 11304–11325.
9. Koslowsky,D.J., Sun,Y., Hindenach,J., Theisen,T. and Lucas,J. (2013) The insect-phase gRNA transcriptome in *Trypanosoma brucei*. *Nucleic Acids Res.*, **42**, 1873–1886.
10. Kirby,L.E., Sun,Y., Judah,D., Nowak,S. and Koslowsky,D. (2016) Analysis of the *Trypanosoma brucei* EATRO 164 bloodstream guide RNA transcriptome. *PLoS Negl. Trop. Dis.*, **10**, e0004793.
11. Read,L.K., Lukes,J. and Hashimi,H. (2016) Trypanosome RNA editing: the complexity of getting U in and taking U out. *Wiley Interdiscip. Rev. RNA*, **7**, 33–51.
12. Zimmer,S.L., Simpson,R.M. and Read,L.K. (2018) High throughput sequencing revolution reveals conserved fundamentals of U-indel editing. *Wiley Interdiscip. Rev. RNA*, **9**, e1487.
13. Sabatini,R. and Hajduk,S.L. (1995) RNA ligase and its involvement in guide RNA/mRNA chimera formation. Evidence for a cleavage–ligation mechanism of *Trypanosoma brucei* mRNA editing. *J. Biol. Chem.*, **270**, 7233–7240.
14. Schmid,B., Read,L.K., Stuart,K. and Goringer,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.
15. Vickerman,K. (1985) Developmental cycles and biology of pathogenic trypanosomes. *Br. Med. Bull.*, **41**, 105–114.
16. Surve,S.V., Jensen,B.C., Heestand,M., Mazet,M., Smith,T.K., Bringaud,F., Parsons,M. and Schnaufer,A. (2017) NADH dehydrogenase of *Trypanosoma brucei* is important for efficient acetate production in bloodstream forms. *Mol. Biochem. Parasitol.*, **211**, 57–61.
17. Zikova,A., Verner,Z., Nenarokova,A., Michels,P.A.M. and Lukes,J. (2017) A paradigm shift: the mitoproteomes of procyclic and bloodstream *Trypanosoma brucei* are comparably complex. *PLoS Pathog.*, **13**, e1006679.
18. Cruz-Reyes,J., Mooers,B.H.M., Doharey,P.K., Meehan,J. and Gulati,S. (2018) Dynamic RNA holo-editosomes with subcomplex variants: insights into the control of trypanosome editing. *Wiley Interdiscip. Rev. RNA*, **9**, e1502.
19. Aphasizheva,I., Alfonzo,J., Carnes,J., Cestari,I., Cruz-Reyes,J., Goringer,H.U., Hajduk,S., Lukes,J., Madison-Antenucci,S., Maslov,D.A. *et al.* (2020) Lexis and grammar of mitochondrial RNA processing in trypanosomes. *Trends Parasitol.*, **36**, 337–355.
20. Madina,B.R., Kumar,V., Metz,R., Mooers,B.H., Bundschuh,R. and Cruz-Reyes,J. (2014) Native mitochondrial RNA-binding complexes in kinetoplastid RNA editing differ in guide RNA composition. *RNA*, **20**, 1142–1152.
21. Kumar,V., Madina,B.R., Gulati,S., Vashisht,A.A., Kanyumbu,C., Pieters,B., Shakir,A., Wohlschlegel,J.A., Read,L.K., Mooers,B.H. *et al.* (2016) REH2C helicase and GRBC subcomplexes may base pair through mRNA and small guide RNA in kinetoplastid editosomes. *J. Biol. Chem.*, **291**, 5753–5764.
22. Aphasizheva,I., Zhang,L., Wang,X., Kaake,R.M., Huang,L., Monti,S. and Aphasizhev,R. (2014) RNA binding and core complexes constitute the U-insertion/deletion editosome. *Mol. Cell. Biol.*, **34**, 4329–4342.
23. Dixit,S., Muller-McNicoll,M., David,V., Zarnack,K., Ule,J., Hashimi,H. and Lukes,J. (2017) Differential binding of mitochondrial transcripts by MRB8170 and MRB4160 regulates distinct editing fates of mitochondrial mRNA in trypanosomes. *MBio*, **8**, e02288-16.
24. McAdams,N.M., Simpson,R.M., Chen,R., Sun,Y. and Read,L.K. (2018) MRB7260 is essential for productive protein–RNA interactions within the RNA editing substrate binding complex during trypanosome RNA editing. *RNA*, **24**, 540–556.
25. 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.
26. Stuart,K., Panigrahi,A.K., Schnaufer,A., Drozdz,M., Clayton,C. and Salavati,R. (2002) Composition of the editing complex of *Trypanosoma brucei*. *Phil. Trans. R. Soc. B: Biol. Sci.*, **357**, 71–79.
27. 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.
28. 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.
29. Panigrahi,A.K., Zikova,A., Dalley,R.A., Acestor,N., Ogata,Y., Anupama,A., Myler,P.J. and Stuart,K.D. (2008) Mitochondrial complexes in *Trypanosoma brucei*: a novel complex and a unique oxidoreductase complex. *Mol. Cell. Proteomics*, **7**, 534–545.
30. Weng,J., Aphasizheva,I., Etheridge,R.D., Huang,L., Wang,X., Falick,A.M. and Aphasizhev,R. (2008) Guide RNA-binding complex from mitochondria of trypanosomatids. *Mol. Cell*, **32**, 198–209.
31. 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.
32. Ammerman,M.L., Downey,K.M., Hashimi,H., Fisk,J.C., Tomasello,D.L., Faktorova,D., Kafkova,L., King,T., Lukes,J. and Read,L.K. (2012) Architecture of the trypanosome RNA editing accessory complex, MRB1. *Nucleic Acids Res.*, **40**, 5637–5650.
33. Hashimi,H., Cicova,Z., Novotna,L., Wen,Y.Z. and Lukes,J. (2009) Kinetoplastid guide RNA biogenesis is dependent on subunits of the mitochondrial RNA binding complex 1 and mitochondrial RNA polymerase. *RNA*, **15**, 588–599.
34. Hernandez,A., Madina,B.R., Ro,K., Wohlschlegel,J.A., Willard,B., Kinter,M.T. and Cruz-Reyes,J. (2010) REH2 RNA helicase in kinetoplastid mitochondria: ribonucleoprotein complexes and essential motifs for unwinding and guide RNA (gRNA) binding. *J. Biol. Chem.*, **285**, 1220–1228.
35. Madina,B.R., Kumar,V., Mooers,B.H. and Cruz-Reyes,J. (2015) Native variants of the MRB1 complex exhibit specialized functions in kinetoplastid RNA editing. *PLoS One*, **10**, e0123441.
36. Kumar,V., Doharey,P.K., Gulati,S., Meehan,J., Martinez,M.G., Hughes,K., Mooers,B.H.M. and Cruz-Reyes,J. (2019) Protein features for assembly of the RNA editing helicase 2 subcomplex (REH2C) in *Trypanosome* holo-editosomes. *PLoS One*, **14**, e0211525.
37. Kumar,V., Ivens,A., Goodall,Z., Meehan,J., Doharey,P.K., Hillhouse,A., Hurtado,D.O., Cai,J.J., Zhang,X., Schnaufer,A. *et al.* (2020) Site-specific and mRNA-specific control of accurate mRNA editing by a helicase complex in trypanosomes. *RNA*, **26**, 1862–1881.
38. Lukes,J., Kaur,B. and Speijer,D. (2021) RNA editing in mitochondria and plastids: weird and widespread. *Trends Genet.*, **37**, 99–102.
39. Jarmoskaite,I. and Russell,R. (2014) RNA helicase proteins as chaperones and remodelers. *Annu. Rev. Biochem.*, **83**, 697–725.
40. Gerasimov,E.S., Gasparyan,A.A., Afonin,D.A., Zimmer,S.L., Kraeva,N., Lukes,J., Yurchenko,V. and Kolesnikov,A. (2021) Complete minicircle genome of *Leptomonas pyrrhocoris* reveals sources of its non-canonical mitochondrial RNA editing events. *Nucleic Acids Res.*, **49**, 3354–3370.
41. 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.
42. Simpson,R.M., Bruno,A.E., Bard,J.E., Buck,M.J. and Read,L.K. (2016) High-throughput sequencing of partially edited trypanosome mRNAs reveals barriers to editing progression and evidence for alternative editing. *RNA*, **22**, 677–695.
43. Kirby,L.E. and Koslowsky,D.J. (2017) Mitochondrial dual-coding genes in *Trypanosoma brucei*. *PLoS Negl. Trop. Dis.*, **11**, e0005989.

44. Smith,J.T. Jr, Dolezelova,E., Tylec,B., Bard,J.E., Chen,R., Sun,Y., Zikova,A. and Read,L.K. (2020) Developmental regulation of edited CYb and COIII mitochondrial mRNAs is achieved by distinct mechanisms in *Trypanosoma brucei*. *Nucleic Acids Res.*, **48**, 8704–8723.
45. Zubradt,M., Gupta,P., Persad,S., Lambowitz,A.M., Weissman,J.S. and Rouskin,S. (2017) DMS-MaPseq for genome-wide or targeted RNA structure probing in vivo. *Nat. Methods*, **14**, 75–82.
46. Livak,K.J. and Schmittgen,T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2(-\Delta\Delta C_T)$ method. *Methods*, **25**, 402–408.
47. Brenndorfer,M. and Boshart,M. (2010) Selection of reference genes for mRNA quantification in *Trypanosoma brucei*. *Mol. Biochem. Parasitol.*, **172**, 52–55.
48. McDermott,S.M., Guo,X., Carnes,J. and Stuart,K. (2015) Differential editosome protein function between life cycle stages of *Trypanosoma brucei*. *J. Biol. Chem.*, **290**, 24914–24931.
49. Tomezsko,P.J., Corbin,V.D.A., Gupta,P., Swaminathan,H., Glasgow,M., Persad,S., Edwards,M.D., McIntosh,L., Papenfuss,A.T., Emery,A. *et al.* (2020) Determination of RNA structural diversity and its role in HIV-1 RNA splicing. *Nature*, **582**, 438–442.
50. Reuter,J.S. and Mathews,D.H. (2010) RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinform.*, **11**, 129.
51. Darty,K., Denise,A. and Ponty,Y. (2009) VARNA: interactive drawing and editing of the RNA secondary structure. *Bioinformatics*, **25**, 1974–1975.
52. Wong,R.G., Kazane,K., Maslov,D.A., Rogers,K., Aphasizhev,R. and Simpson,L. (2015) U-insertion/deletion RNA editing multiprotein complexes and mitochondrial ribosomes in *Leishmania tarentolae* are located in antipodal nodes adjacent to the kinetoplast DNA. *Mitochondrion*, **25**, 76–86.
53. Seiwert,S.D. and Stuart,K. (1994) RNA editing: transfer of genetic information from gRNA to precursor mRNA in vitro. *Science*, **266**, 114–117.
54. 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.
55. Mesitov,M.V., Yu,T., Suematsu,T., Sement,F.M., Zhang,L., Yu,C., Huang,L. and Aphasizheva,I. (2019) Pentatricopeptide repeat poly(A) binding protein KPAF4 stabilizes mitochondrial mRNAs in *Trypanosoma brucei*. *Nat. Commun.*, **10**, 146.
56. Tylec,B.L., Simpson,R.M., Kirby,L.E., Chen,R., Sun,Y., Koslowsky,D.J. and Read,L.K. (2019) Intrinsic and regulated properties of minimally edited trypanosome mRNAs. *Nucleic Acids Res.*, **47**, 3640–3657.
57. Medley,J.C., Panzade,G. and Zinovyeva,A.Y. (2021) microRNA strand selection: unwinding the rules. *Wiley Interdiscip. Rev. RNA*, **12**, e1627.
58. Gerasimov,E.S., Afonin,D.A., Korzhavina,O.A., Lukes,J., Low,R., Hall,N., Tyler,K., Yurchenko,V. and Zimmer,S.L. (2022) Mitochondrial RNA editing in *Trypanoplasma borreli*: new tools, new revelations. *Comput. Struct. Biotechnol. J.*, **20**, 6388–6402.
59. 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.
60. 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.
61. Li,F., Herrera,J., Zhou,S., Maslov,D.A. and Simpson,L. (2011) Trypanosome REH1 is an RNA helicase involved with the 3'–5' polarity of multiple gRNA-guided uridine insertion/deletion RNA editing. *Proc. Natl Acad. Sci. USA*, **108**, 3542–3547.
62. Missel,A., Souza,A.E., Norskau,G. and Goringer,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.
63. Cruz-Reyes,J., Mooers,B.H., Abu-Adas,Z., Kumar,V. and Gulati,S. (2016) DEAH-RHA helicase•Znf cofactor systems in kinetoplastid RNA editing and evolutionarily distant RNA processes. *RNA Dis.*, **3**, e1336.
64. Kruse,E., Voigt,C., Leeder,W.M. and Goringer,H.U. (2013) RNA helicases involved in U-insertion/deletion-type RNA editing. *Biochim. Biophys. Acta*, **1829**, 835–841.
65. Stoltzfus,A. (1999) On the possibility of constructive neutral evolution. *J. Mol. Evol.*, **49**, 169–181.
66. Gray,M.W., Lukes,J., Archibald,J.M., Keeling,P.J. and Doolittle,W.F. (2010) Cell biology. Irremediable complexity?*Science*, **330**, 920–921.
67. Lukes,J., Archibald,J.M., Keeling,P.J., Doolittle,W.F. and Gray,M.W. (2011) How a neutral evolutionary ratchet can build cellular complexity. *IUBMB Life*, **63**, 528–537.
68. Schnaufer,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.
69. Hierro-Yap,C., Subrtova,K., Gahura,O., Panicucci,B., Dewar,C., Chinopoulos,C., Schnaufer,A. and Zikova,A. (2021) Bioenergetic consequences of F_oF_1-ATP synthase/ATPase deficiency in two life cycle stages of *Trypanosoma brucei*. *J. Biol. Chem.*, **296**, 100357.
70. Schnaufer,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.
71. Dean,S., Gould,M.K., Dewar,C.E. and Schnaufer,A.C. (2013) Single point mutations in ATP synthase compensate for mitochondrial genome loss in trypanosomes. *Proc. Natl Acad. Sci. USA*, **110**, 14741–14746.
72. 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.
73. 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.