CellPress

Review

U-Insertion/Deletion mRNA-Editing Holoenzyme: Definition in Sight

Inna Aphasizheva[1,*] and Ruslan Aphasizhev[1,2]

RNA editing is a process that alters DNA-encoded sequences and is distinct from splicing, 5′ capping, and 3′ additions. In 30 years since editing was discovered in mitochondria of trypanosomes, several functionally and evolutionarily unrelated mechanisms have been described in eukaryotes, archaea, and viruses. Editing events are predominantly post-transcriptional and include nucleoside insertions and deletions, and base substitutions and modifications. Here, we review the mechanism of uridine insertion/deletion mRNA editing in kinetoplastid protists typified by *Trypanosoma brucei.* This type of editing corrects frameshifts, introduces translation punctuation signals, and often adds hundreds of uridines to create protein-coding sequences. We focus on protein complexes responsible for editing reactions and their interactions with other elements of the mitochondrial gene expression pathway.

Is There an Adaptive Advantage in Complexity?

In 1986, Rob Benne and co-workers described the insertion of four uridines into cytochrome *c* oxidase subunit 2 (CO2) mRNA from *T. brucei* as means of correcting the encoded frameshift at the RNA level [1]. Astutely named RNA editing, this phenomenon later came to symbolize massive U insertions [2] and U deletions [3] that create open reading frames in transcripts of cryptic mitochondrial genes in kinetoplastid protists. This paradigm-shifting discovery stimulated researchers to look closer at discrepancies between DNA and RNA sequences in other organisms and ultimately led to identification of several divergent and largely unrelated editing mechanisms, such as A to I [4] and C to U base deamination [5], 3′-to-5′ polymerization [6], and others. The narrow phylogenetic distribution of editing systems suggests their derived character within lineages in which they currently exist rather than editing being a primordial trait retained from a common evolutionary ancestor in some organisms and lost in others [7]. The sheer mechanistic and component complexity, and the lack of apparent adaptive advantage of having one, positions trypanosomal editing as a fruitful platform for evolutionary debate on the origins of macromolecular assemblies. The constructive neutral evolution (CNE) hypothesis argues that the functional editing machinery may evolve in the absence of positive selection and, importantly, prior to the actual need for the editing process [8]. By virtue of extant proteins forming neutral mutation-driven interactions, for example, an enzyme with an RNA-binding protein, such assembly may acquire a novel capacity. Without pressure from purifying selection, the neutral capacity can persist with no essential cellular function until a mutation arises that can be corrected by such pre-existing activity. Thus, the detrimental impact of a gene mutation may not be compensated unless a functional system to correct the sequence at the RNA level is already in place. It follows that editing is an intrinsically mutagenic process: once evolved, the editing system allows accumulation of mutations that otherwise would be eliminated by the purifying selection [9,10]. Likewise, accumulation of multiple mutations would make reverse

Trends

Uridine insertion/deletion editing generates protein-coding sequences in most mitochondrial mRNAs of trypanosomes. The emerging architecture of the editing holoenzyme suggests an RNA-mediated assembly of the multisubunit enzymatic RNA editing core and RNA editing substrate binding (RESC) complexes.

Recently characterized RESC complex is composed of ~17 polypeptides that can be clustered into guide RNA (gRNA) binding complex, RNA editing mediator complex, and polyadenylation mediator complex modules. These modules are responsible for gRNA binding, and mediating interactions with the enzymatic core editing and polyadenylation complexes, respectively.

The majority of RNA editing factors are essential for parasite viability and do not have apparent human homologs. Therefore, RNA editing pathway represents a significant source of therapeutic targets relevant to neglected tropical diseases.

[1]Department of Molecular and Cell Biology, Boston University School of Dental Medicine, Boston, MA 02118, USA
[2]Department of Biochemistry, Boston University School of Medicine, Boston, MA 02118, USA

*Correspondence: innaaf@bu.edu (I. Aphasizheva).

 http://dx.doi.org/10.1016/j.pt.2015.10.004

changes all but impossible, and render editing an essential pathway. In this context, the composition of the enzymatic core editing complex proved most instructive: catalytic modules implicated in fundamental cellular functions, such as DNA repair and RNA interference (RNAi), along with proteins likely acquired by horizontal gene transfer, operate as stable protein complex that cleaves mRNA, adds or removes Us, and re-ligates fragments. Finally, when editing becomes an indispensable process, such as generation of a protein-coding sequence, it must be incorporated into the overall gene expression pathway. It could be expected that interactions of the editing machinery with RNA processing and translation complexes would be as unique as editing systems themselves. Selection of correctly edited mRNA by the ribosome in a background of **partially edited** and unedited transcripts is among the most obvious problems that require additional levels of control. These considerations do not rule out possible adaptive advantage of editing once it evolved–indeed two reports indicate that alternative editing may generate protein diversity [11,12]. Editing-dependent protein diversity, both the fact and the function, remain to be firmly established, leaving the question wide open to future investigation and hypothesis building. Here, we review the complexity of trypanosomal insertion/deletion editing in terms of underlying biochemistry and potential origins of editing effectors, as well as determinants that direct position-specific insertion and deletion of uridines.

Elemental Editing Reactions Are Catalyzed by Modular RNA Editing Core Complex (RECC)

T. brucei, the causative agent of African sleeping sickness, and most other representatives of **Kinetoplastea**, such as *Leishmania* spp., are characterized by the presence of the **kinetoplast**. This disc-shaped, high-density nucleoprotein structure is located in the mitochondrial lumen adjacent to the flagellar base. The kinetoplast encloses the mitochondrial genome (**kinetoplast DNA, kDNA**), which is composed of two types of catenated circles. Relatively few **maxicircles** (~25 kb) encode genes typically found in mitochondrial genomes, such as rRNAs, ribosomal protein RPS12 and subunits of respiratory complexes, while thousands of ~1-kb **minicircles** constitute the bulk of kDNA. In *T. brucei*, six of the 18 annotated mRNAs encode predicted polypeptides, while the remaining 12 transcripts must undergo editing to acquire open reading frames and translation punctuation signals. The product of trypanosomal mRNA editing is not collinear with DNA as it contains extra nucleotides compared to the gene sequence, and sometime lacks encoded uridines. Historically, the determinants of position-specific U insertions and deletions have been discovered as short patches of complementarity between edited mRNA and maxicircle DNA in *Leishmania tarentolae* [13]. By allowing for wobble G-U, in addition to canonical Watson–Crick base-pairing, short [50–60 nucleotides (nt)] mitochondrial RNAs transcribed from maxicircles have been recognized as carriers of genetic information and termed **guide RNAs (gRNAs)**. Further work established that most gRNAs are encoded in minicircles [14]. The predicted secondary structure of gRNA–mRNA hybrid instantly suggested a mechanism by which the editing site and the extent of U insertions/deletions are determined without invoking template-dependent polymerization of nucleic acids [13]. The initial site selection is accomplished via a short (5–10-nt) region of complementarity between the **5′ anchor** of gRNA and **pre-edited mRNA**. The rest of gRNA forms an imperfect duplex (**3′ anchor**) with mRNA, which results in bulging of single-stranded uridines in mRNA (deletion sites) or purine nucleotides in gRNA (insertion sites, Figure 1). At either site, the mRNA is cleaved at the first unpaired nucleotide adjacent to the 5′ anchor duplex. The resultant deletion and insertion intermediates are distinct: single-stranded uridines become exposed to a 3′–5′ exonucleolytic attack in the former, while a gap is created in the latter. Upon trimming uridines to the first paired base in the deletion site, or adding gRNA-specified numbers of uridines to a **5′ cleavage fragment** in the insertion site, the **5′** and **3′ cleavage fragments** are joined to restore mRNA continuity, which extends the double-stranded anchor region.

Glossary

3′ anchor: refers to scattered base-pairing between gRNA region that directs editing reactions and pre-edited mRNA.
5′ anchor: 5′ part of the gRNA that forms a continuous 5–10-nt duplex with pre-edited mRNA; this region is responsible for initial gRNA–mRNA interaction.
5′ and 3′ cleavage fragments: mRNA fragments generated by gRNA-directed endonucleolytic cleavage; the first reaction of the editing cascade.
dsRNA: double-stranded RNA.
Endo/exo/phosphatase (EEP): metal-dependent hydrolase, endonuclease/exonuclease/phosphatase family.
Editing block: mRNA segment covered by a single gRNA; typically contains both U-insertion and U-deletion sites.
Editing domain: mRNA segment covered by multiple overlapping gRNAs. Sequence changes directed by the initiating gRNA create binding site for a sequential one and so forth; the hierarchical gRNA binding provides for the overall 3′–5′ progression of editing events within the domain.
Editing site: position of the guide RNA-directed mRNA cleavage where uridines are either removed from or added to the 5′ cleavage fragment.
Fully edited mRNA: final product of the editing process; contains an open reading frame.
Guide RNA: small non-coding RNA that specifies positions and extent of uridine insertions and deletions. gRNAs are typically 40–60 nt in length and possess 5′ triphosphates and 3′ oligo U-tails.
Kinetoplast: a densely packed, disc-like nucleoprotein structure that encloses mitochondrial DNA.
Kinetoplast DNA (kDNA): mitochondrial genome, DNA component of the kinetoplast composed of maxicircles and minicircles. In most cases circular DNA molecules are locked into a catenated network, but in some species circles are not catenated and distributed throughout mitochondrial lumen.
Kinetoplastida (Kinetoplastea): a group of flagellated protists belonging to the phylum Euglenozoa and characterized by the presence of a kinetoplast.