Journal of Bioenergetics and Biomembranes, Vol. 26, No. 2, 1994

RNA Editing in Mitochondria of Cultured Trypanosomatids: Translatable mRNAs for NADH-dehydrogenase Subunits Are Missing

Paul Sloof,[1] Gert Jan Arts,[1] Janny v. d. Burg,[1] Hans v. d. Spek,[1] and Rob Benne[1]

Received October 15, 1993; accepted December 12, 1993

RNA editing in mitochondria of kinetoplastid protozoa involves the posttranscriptional insertion and deletion of uridylate residues in protein encoding regions of pre-mRNAs. Editing is required to remove gene-encoded translational defects or to convert a nonsense sequence into a sense message. In cultured trypanosomatids, however, translationally defective pre-mRNAs for a number of NADH-dehydrogenase subunits are not converted into functional mRNAs by editing. In this report, the available data are discussed in the context of current models for RNA editing.

KEY WORDS: Trypanosomatidae; kinetoplast; mitochondrion; RNA editing; guide RNA.

INTRODUCTION

Characteristics of the Editing Process

RNA editing in kinetoplastid protozoa is an unusual form of RNA processing (Benne *et al.*, 1986) that is required for the expression of a special class of mitochondrial (mt) genes, the so-called cryptogenes (Simpson and Shaw, 1989), which produce nonfunctional precursor (pre-)mRNAs. The conversion of the pre-mRNAs into functional mRNAs involves the addition and occasional deletion of uridylate residues. In some transcripts editing is limited to a small region in which it generates translational start codons or corrects gene-encoded frame shifts. In other RNAs, large sections are altered by extensive editing (pan-editing), creating a functional message in the process. In the following sections we will briefly summarize the available data on the biochemistry of the editing process and discuss some of the current models; for more detailed information, we refer to Benne, 1993; Sloof and Benne, 1993; Stuart, 1993a,b; Hajduk *et al.*, 1993; Simpson *et al.*, 1993; Simpson and Shaw, 1989.

The editing information is derived from small, mitochondrially encoded guide (g)RNAs which are partially complementary to edited transcript sequences, if G:U base pairing is allowed in addition to conventional Watson/Crick pairs (Blum *et al.*, 1990). Inspection of gRNA sequences revealed three domains which may have a role at different stages of the editing process (Blum *et al.*, 1990; Blum and Simpson, 1990; see Fig. 1): (i) a 4 to 18 nucleotides long anchor sequence in the 5′ region of the gRNA, which is complementary to a target region in pre-mRNA that is located 3′ of an editing region (defined as the pre-mRNA region to be edited by a single gRNA). The formation of a duplex between these two sequences most likely initiates the editing process. (ii) An informational domain at the 3′ flank of the anchor domain, which is complementary to a section of edited mRNA. This domain is proposed to guide the U-sequence alteration process by extension of the base-paired region. (iii) A nonencoded oligo-U extension at the 3′ terminus of the gRNA.

The U-sequence alteration processes may be carried out in a series of two-step reactions (Blum *et*

[1] E. C. Slater Institute, Biocentre Amsterdam, Academic Medical Centre, Meibergdreef 15, 1105 AZ Amsterdam, The Netherlands.

0145-479X/94/0400-0193$07.00/0

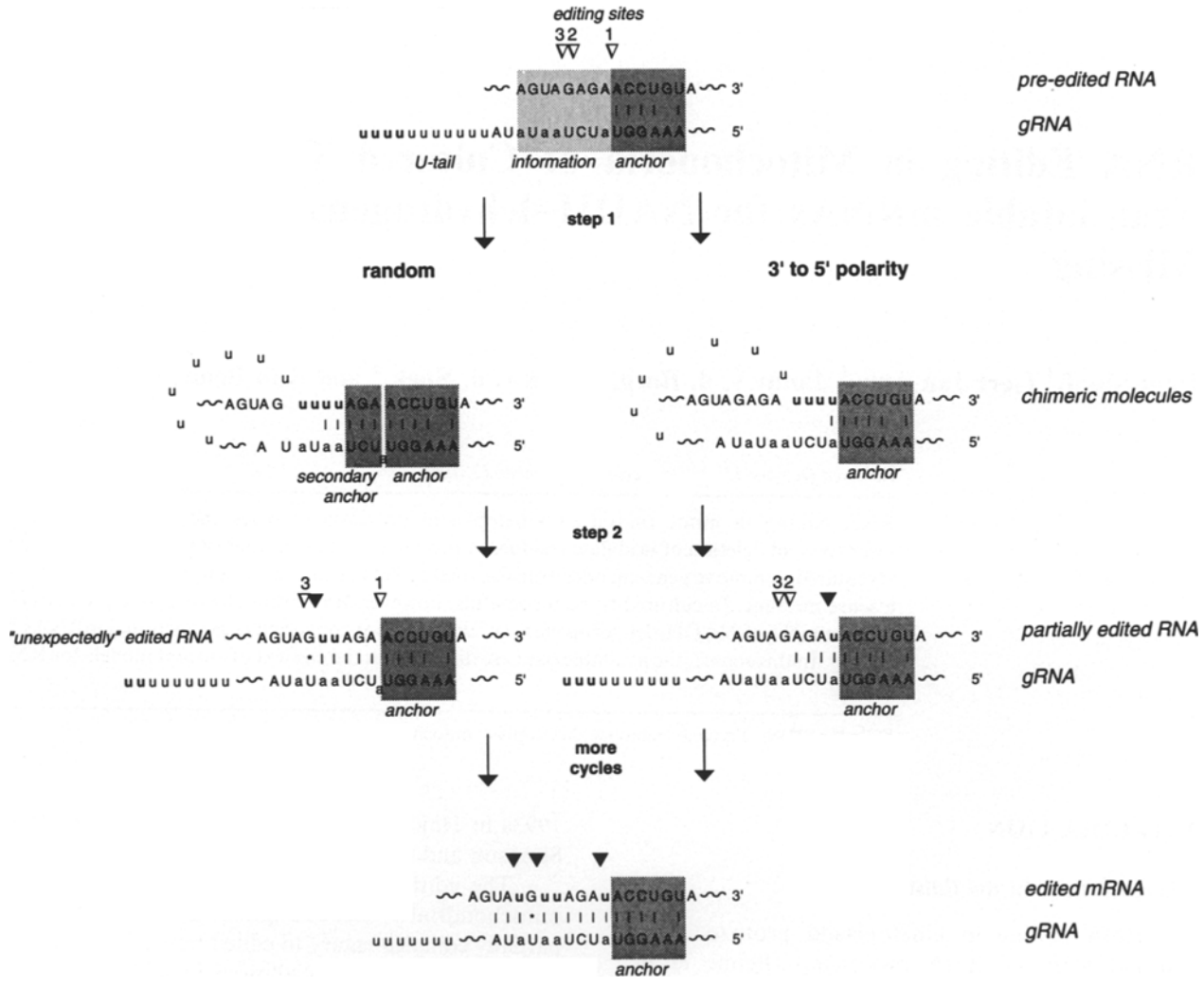

Fig. 1. gRNA : pre-mRNA interactions in RNA editing. This model is adapted from Blum *et al.* (1991). Editing of sites (open arrowheads) in a pre-mRNA (top line) is initiated by formation of a duplex between the anchor domain of a cognate gRNA (bottom line) and a target region in pre-mRNA, 3′ of a region to be edited. A two-step pathway has been proposed for the transfer of (a) U residue(s) from the gRNA U tail to the edited site (black arrowheads) in mRNA, in which either concerted RNA-mediated transesterifications or enzymatic cleavage-ligation reactions are involved: in a first step a chimeric RNA is formed in which a gRNA is attached to an editing site of pre-mRNA via its U tail and the duplexed region is extended by the participation of U tail residues (u), and in a second step the chimeric RNA is resolved, leaving the base-paired U-tail residues behind, between the linked 5′ and 3′ moieties of pre-mRNA. Two possible routes are shown: On the right panel editing occurs one U at a time with a strict 3′ to 5′ polarity (starting at site 1 in the figure); on the left panel the order of editing is random (starting at site 2 in the figure). Resolution of chimeric RNAs leads to a partially edited RNA in the right-hand route, or to an unexpectedly edited RNA in the left-hand route. In both routes, the opening of the pre-mRNA and the resolution of the chimeric RNA occurs at the phosphodiester bond 3′ of the first mismatched residue 5′ from the anchor duplex (editing site 1 on the right panel) or 5′ from a secondary anchor duplex (editing site 2 on the left panel). More cycles lead to a completely edited mRNA. For deletion of U residues from pre-mRNA in editing, a 3′ U exonuclease step has been invoked to remove the U's from the 5′ moiety of pre-mRNA generated by step 1 (Blum *et al.*, 1991). For further details, see text.

al., 1991; see Fig. 1). In step 1, the pre-mRNA moiety of the gRNA:pre-mRNA duplex is cleaved at the phosphodiester bond located 3′ of the first mismatching residue. A chimeric molecule is formed by connecting a gRNA via its U-tail to the 3′ moiety of the pre-mRNA. The duplexed region is extended by the participation of gRNA U-tail residues. In step 2 the chimeric RNA is resolved by cleavage of a phosphodiester linkage similarly located 3′ of the first mismatching residue, and a partially edited RNA is formed by joining the 5′ and 3′ moieties with the U's derived from the gRNA in between. Additional

rounds will result in a completely edited mRNA. This scenario of the editing process is based on the properties of gRNAs as discussed above and on the demonstration of the existence of chimeric RNAs in mtRNA (Blum *et al.*, 1991; Read *et al.*, 1992a; Arts *et al.*, 1993). Similar events have been proposed for U-deletion processes, which for reasons of limited available space will not be discussed here (but see Blum *et al.*, 1991 and Cech, 1991).

The sites of an editing region in pre-mRNA may be edited either with a strict 3′ to 5′ polarity with consecutive editing of sites 1, 2, and 3 (right-hand panel of Fig. 1), or without a preferred order (the random model, e.g., editing of site 2 precedes that of 1 and 3 or vice versa, left-hand panel of Fig. 1). The polarity model is based on the characteristics of a set of partially edited RNAs from *Leishmania tarentolae*, which could represent intermediates of such a strict 3′ to 5′ editing process in which one site at a time is edited with one U at a time (Sturm and Simpson, 1990). An alternative explanation for the presence of partially edited RNAs, however, is that they could be generated by editing with gRNAs containing truncated informational domains, and/or short U-tails. These gRNA forms are abundantly present in mitochondria of *Crithidia fasciculata* (Arts *et al.*, 1993) and they participate in chimer formation in a number of species (Arts *et al.*, 1993; Read *et al.*, 1992a; Blum *et al.*, 1991). The occasionally abundant presence of so-called unexpectedly edited sequences in which 5′ sites are edited before (complete) editing of 3′ sites and in which nonediting sites can also be edited (Sturm and Simpson, 1990; Sturm *et al.*, 1992; Decker and Sollner-Webb, 1990; Koslowsky *et al.*, 1991) could therefore be indicative of a more random order of events during the editing process. It has been pointed out, however, that some of the unexpected sequences are created by editing with a noncognate gRNA ("misediting," Maslov and Simpson, 1992; for a more extensive discussion of the polarity of editing, see Maslov and Simpson, 1992; Benne, 1993; and Sloof and Benne, 1993). A major problem facing those attempting to assess the relevance of the various partially and unexpectedly edited transcripts and that of the different chimeric RNAs is the lack of an efficient *in vitro* editing system in which their role can be tested. The importance of the anchor sequence for chimer formation has been demonstrated, nevertheless, in *in vitro* systems (Blum and Simpson, 1992; Arts *et al.*, unpublished), providing evidence in support of at least the first part of the models.

The mechanisms by which phosphodiester linkages are broken and resealed in editing is unknown, but it has been invoked that they involve either a series of transesterifications, reminiscent of those occurring in (self)splicing, or a cascade of enzymatic "cut and paste" reactions, in which (a) pre-mRNA editing site-specific endonuclease(s) and an RNA ligase could participate (Cech, 1991; Blum *et al.*, 1991). In either case, gRNA U-tails are (re)charged by a terminal uridylyl transferase (TUTase) activity. Although at present a definite choice cannot be made, the current evidence slightly favors the enzyme cascade model: (i) the enzymatic activities required have been found in trypanosomatid mitochondria (Bakalara *et al.*, 1989; Harris *et al.*, 1992; Simpson *et al.*, 1992), (ii) both RNA ligase and TUTase co-sediment with gRNA and pre-mRNA in high-molecular-weight complexes during glycerol gradient centrifugation of mitochondrial lysates (Pollard *et al.*, 1992), and (iii) chimeric molecules have been generated from synthetic gRNA and pre-mRNA *in vitro* with heterologous RNA ligase and endonuclease (Sollner-Webb, 1992).

The order in which gRNAs act in the editing of an extensively edited RNA can be deduced from the characteristics of complete sets of gRNAs that have been identified for panediting of transcripts for small subunit ribosomal protein (RP) S12 and ATPase subunit 6 (also known as maxicircle unidentified reading frame (MURF) 4; see legend to Fig. 1) in *L. tarentolae* (Maslov and Simpson, 1992). Only the target sequence for the anchoring of the gRNA that edits the most 3′ region of an editing domain is present in unedited pre-mRNA. The target sequences for anchoring of the other gRNAs are created via the editing by their immediate predecessors. Most RNAs consist of one such editing domain; *L. tarentolae* RPS12, however, has three independently edited domains (Maslov and Simpson, 1992), and *Trypanosoma brucei* NADH-dehydrogenase (ND) subunit 7 RNA has two such domains (Koslowsky *et al.*, 1990). The 3′ to 5′ progress of editing is secured by the skewed distribution of G:U pairs in gRNA–mRNA duplexes which predominate in the informational domain and are scarce in the anchor duplex. As a result, the anchor of the incoming gRNA can form a more stable interaction with edited RNA, thereby replacing the preceding gRNA (Maslov and Simpson, 1992). This scenario is in agreement with the characteristics of partially edited cytochrome oxidase subunit (cox 3) RNAs and other partially (pan)edited

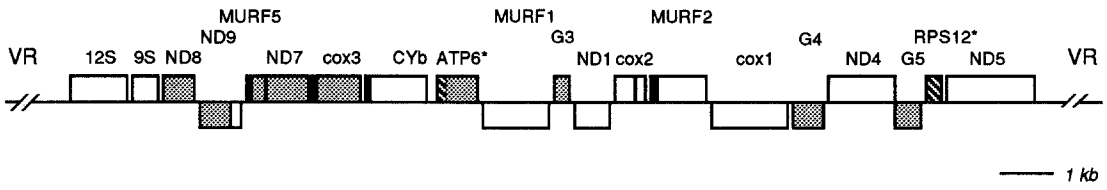

Fig. 2. Gene organization in the maxicircle of *T. brucei*, *L. tarentolae*, and *C. fasciculata*. The composite gene map of the maxicircle of the three trypanosomatid species indicates the genes for the 12S and 9S ribosomal RNAs; cytochrome *c* oxidase (cox) subunits 1, 2, and 3; NADH dehydrogenase (ND) subunits 1, 4, 5, 7, 8, and 9; apocytochrome *b* (CYb); small subunit ribosomal protein (RP) S 12; ATPase (ATP) subunit 6; maxicircle unidentified reading frame (MURF) 1, 2, and 5; and transcribed G-rich regions (G) 3, 4, and 5. The asterisks indicate that the assignment of the ATPase 6 (otherwise called MURF4) and RPS 12 (G6) genes is controversial (Simpson *et al.*, 1993; Stuart, 1993b; Maslov *et al.*, 1992; Souza *et al.*, 1992; Read *et al.*, 1992; Bhat *et al.*, 1990). Regions that correspond to RNA (section)s in which editing has been demonstrated to occur in all species are indicated in black, those for which only the data for *T. brucei* and *L. tarentolae* are available are hatched. Cryptogenes for RNAs which are pan-edited in *T. brucei* exclusively are in gray. Unedited RNA(section)s are indicated by the white boxes. The direction of transcription is from left to right for the genes above the line; the genes below the line are transcribed in the opposite direction. VR indicates the variable region containing repetitive DNA sequences (Sloof *et al.*, 1992).

RNAs in which the edited sections are invariably found 3′ of the unedited parts (Feagin *et al.*, 1988a; Abraham *et al.*, 1988; Stuart, 1993a,b).

Genomic Organization in Trypanosome mt DNA

The mitochondrial genome of the Trypanosomatidae is a complex network of catenated maxicircles ad minicircles (reviewed in Borst and Hoeijmakers, 1979; Simpson, 1987). The maxicircle component harbors genes for 18 putative proteins, only four of which are conventional in the sense that their primary transcripts are not edited and are translatable as such (reviewed in Sloof and Benne, 1993; Stuart 1993a,b; Simpson *et al.*, 1993; see Fig. 2). All other protein genes are cryptic, producing untranslatable or nonsense pre-mRNAs that require editing. The extent to which editing occurs varies between different RNAs and also between trypanosome species, however. Editing appears to be most extensive in *T. brucei*, in which nine RNAs are panedited. In RNAs from cultured *L. tarentolae* and *C. fasciculata* the editing seems to be less spectacular, panediting having been observed, so far, in two *L. tarentolae* RNAs only (see next section).

gRNA genes have been identified both in minicircles and maxicircles (for reviews see Stuart, 1993a,b; Simpson *et al.*, 1993; Benne, 1993; Sloof and Benne, 1993). To date, about 50 gRNA genes have been found in *T. brucei* (Pollard *et al.*, 1990; Corell *et al.*, 1993), only three of which have a maxicircle location (Van der Spek *et al.*, 1991). It is estimated that the extensive editing observed in this organism would require about 240 different gRNAs (Stuart, 1993b). Since each minicircle harbors at least three gRNA genes (Stuart, 1993b) and since about 250 different sequence classes exist in this species (Steinert and Van Assel, 1980), there appears to be more than sufficient coding capacity. In fact, many of the gRNAs specifying the editing of one RNA display substantial overlap, suggesting some form of proofreading (Pollard *et al.*, 1990; Corell *et al.*, 1993). In *L. tarentolae* nine maxicircle and 17 minicircle encoded gRNAs have been identified, 23 of which are sufficient to account for all the editing observed in this organism (Blum *et al.*, 1990; Sturm and Simpson, 1991; Maslov and Simpson, 1992).

NADH Dehydrogenase Genes Produce Translationally Defective RNAs in Cultured Trypanosomatids

In *T. brucei*, the editing of most transcripts is developmentally regulated. For example, the apocytochrome *b* (CYb) and cytochrome oxidase subunit (cox) 2 transcripts are edited in cultured, procyclic forms (PF) but remain essentially unedited, and thus untranslatable, in bloodstream forms (BF) (Feagin *et al.*, 1987; Feagin and Stuart, 1988). In contrast, another set of transcripts including RPS12 RNA, G3 RNA (G for G-rich), G4 RNA, and RNAs encoding ND7, 8, and 9 are panedited into translatable RNAs exclusively in BF but not in PF *T. brucei* (Read *et al.*, 1992b; Stuart, 1993a,b; Koslowsky *et al.*, 1990; Souza *et al.*, 1992). Only cox3 RNA, ATP6 (= MURF4) RNA, and G5 RNA appear to be edited in both life cycle stages (Feagin *et al.*, 1988a; Bhat *et al.*, 1990; Stuart, 1993a,b). In this organism, RNA editing is clearly used as a control point for gene expression during the different

life-cycle stages. It remains obscure how the differential editing of certain transcripts is brought about, but it can be speculated that the presence or absence of gRNAs plays a decisive role. So far, however, evidence to support such a hypothesis is lacking since both life-cycle stages contain some of the necessary gRNAs (Koslowsky *et al.*, 1992). The analysis is far from complete at present, however, and it remains entirely possible that gRNAs exist that play a key role in this respect.

Not much is known about different life-cycle stages (if any) of the other trypanosomatid species that are frequently used in RNA editing research: the lizard trypanosomatid *L. tarentolae* and the insect parasite *C. fasciculata*. All of the information on the extent of RNA editing in these organisms is derived from forms that have been cultured in the laboratory for extended periods of time (sometimes more than 50 years). As mentioned above, in *L. tarentolae* only two panedited RNAs [ATP6 (= MURF4) and RPS12] have been identified. All of the other G-rich regions appear to produce unedited and therefore untranslatable pre-mRNAs, although unexpectedly edited sequences at the 3′ terminus of these RNAs have been reported (Maslov and Simpson, 1992; Simpson *et al.*, 1993; Souza *et al.*, 1992). In our laboratory we have performed an extensive search for edited ND8 and 9 RNAs in cultured *C. fasciculata*, which included Northern blot analysis, PCR experiments, and the sequencing of numerous cDNAs. In this report we summarize our results which indicate that also in the cultured form of this trypanosomatid ND8 and ND9 RNAs are largely unedited and remain untranslatable. We discuss these data in the context of current models for RNA editing and our previous observations that other ND subunit RNAs in cultured *C. fasciculata* are also untranslatable: ND1 RNA is unedited and lacks therefore a translational initiation codon (Van der Spek *et al.*, 1990) and the editing of ND7 RNA moves a potential translational initiation codon out of frame (Van der Spek *et al.*, 1988).

EXPERIMENTAL

Oligonucleotides

The oligonucleotides used in this study are listed below:

C82: 5′-CGGATTTTATTTGGAGGGAC-3′, coordinates (Sloof *et al.*, 1987) 2233-2252.

C84: 5′-AATAAATTTTAGGTTAAACTT-GATTT-3′, coordinates (Sloof *et al.*, 1987) 2567–2541.

C93: 5′-AATTATATTAAATTTAATAAAA-CACAAC-3′ (see Fig. 3A).

Cell Culture and Nucleic Acid Isolations

C. fasciculata was cultured according to Kleisen *et al.*, 1975. Total and mitochondrial RNA was isolated as described by Van der Spek *et al.*, 1990. The enrichment of mitochondrial RNA for ribosomal 9S and 12S RNAs was 50- to 100-fold as judged from Northern blot analysis.

cDNA Cloning and Sequencing

ND8 and nine cDNAs were isolated after screening a cDNA library, constructed from total RNA of *C. fasciculata* (Van der Spek *et al.*, 1988), using genomic probes encompassing the *C. fasciculata* maxicircle region between positions 2190 and 2555 (for ND8), and 2850 and 2560 (for ND9) (Sloof *et al.*, 1987). Specific cloning of ND8 cDNAs was performed as follows: (i) cDNA was synthesized by reverse transcription of mitochondrial RNA using oligo dT or a primer complementary to ND8 RNA (C84 or C93, see Fig. 3A). (ii) Amplification of cDNAs was performed by PCR using a second primer, specific for the 5′ terminal region of unedited ND8 (primer C82). PCR products were ligated into the polylinker region of PUC19 and cloned in *E. coli* DH 5α. cDNA and PCR procedures have been described previously (Arts *et al.*, 1993). ND8 cDNAs were sequenced according to Sanger *et al.* (1977) with a Pharmacia T7 sequencing kit and standard PUC forward and reverse primers. ND1 and ND7 cDNA analysis has been described previously (Van der Spek *et al.*, 1988, 1990). For ND9 cDNAs, similar procedures were used (not shown).

Electrophoresis of RNA, Northern Blotting, and Hybridization

Electrophoresis of glyoxylated RNA on 1.75% agarose gels and blotting was performed according to Van der Spek *et al.*, (1990). UV crosslinked filters were pre-hydridized for 1 to 2 hours and hybridized for 14–18 hours in hybridization mixture (Sambrook *et al.*, 1989) at 30°C in a minimal volume (6–16 ml), washed for three consecutive periods of 2 min in 6 × SSC, 0.1% SDS at room temperature, elevated

temperature (Itakura *et al.*, 1984), and room temperature. Oligonucleotides were phosphorylated at the 5′ end using T4 polynucleotide kinase (Pharmacia) and [γ-^{32}P]ATP (Amersham).

RESULTS

ND8 and ND9

In most cases in which panedited RNAs were identified, a multi-step selection procedure was applied in which partially edited cDNAs with varying extents of editing were used to find cDNAs derived from completely edited transcripts (for examples, see Feagin *et al.*, 1988a; Koslowsky *et al.*, 1990; Bhat *et al.*, 1990; Souza *et al.*, 1992; Maslov *et al.*, 1992). We also followed this approach in *C. fasciculata*. Extensive cDNA analysis in which more than 50 cDNAs were screened following their selection from a library with an unedited probe resulted in the isolation of many unedited cDNA clones. These exhibited a variable polyadenylation site (Fig. 3C), a feature which has been observed for other trypanosomatid mitochondrial RNAs (Souza *et al.*, 1992; Van der Spek *et al.*, 1990). The analysis also yielded one cDNA, which was partially edited in the 3′ terminal region by the insertion of 16 and the deletion of 3 U's, respectively (clone 45, see Fig. 3A and B). The amino acid sequence encoded by the edited section contains a high percentage of identical and conserved amino acids when compared to that of the C-terminus of the inferred *T. brucei* protein (see Fig. 4A). Moreover, the translational termination codons that are created by editing in both cases are located exactly at the position at which the high degree of similarity between the GAC sequences (i.e., the sequence minus the U's encoded by the gene) of the unedited pre-mRNAs of the two species breaks down (indicated by the arrowhead in Fig. 4B). This most likely indicates that the edited *C. fasciculata* sequences are produced by the correct, cognate gRNA and do not represent a misedited RNA. In the latter case the *C. fasciculata* sequence would most certainly be different from the *T. brucei* sequence and not be located at a corresponding position.

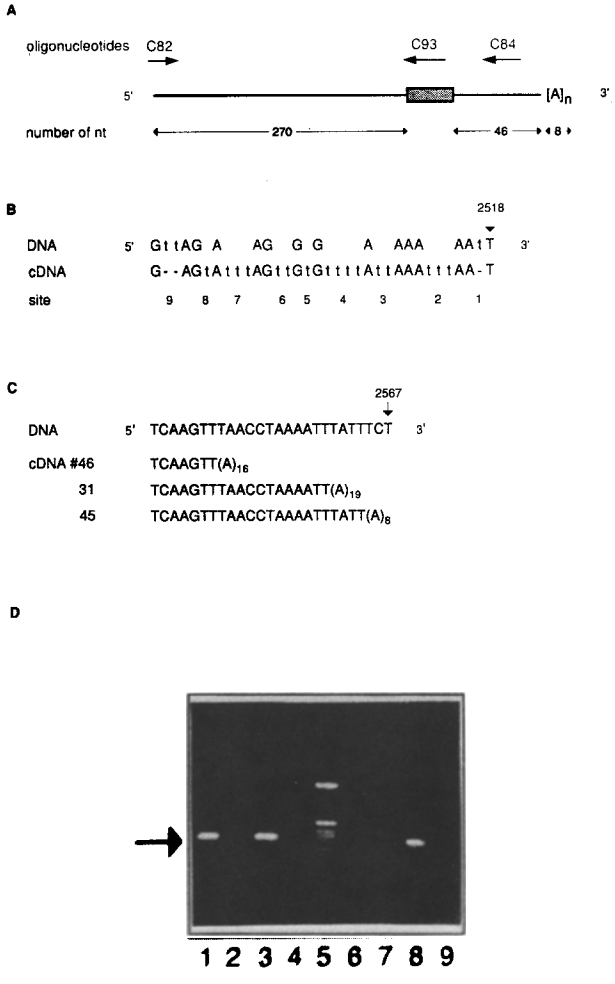

Fig. 3 Analysis of ND8 cDNAs. (A) Schematic representation of cDNA clone 45. The edited domain, indicated by the boxed region, is located 46 nucleotides upstream of the poly(A) tail of 8 residues in length and 270 nucleotides downstream from of the 5′ terminus of the transcript. The transcript has been mapped in the genomic sequence from position 2234 to 2564 (Sloof *et al.*, 1987). The position from 5′ to 3′ of the oligonucleotides C82, C84, and C93 in this cDNA are indicated by arrows (for sequences see the Experimental section). (B) Nucleotide sequence of the edited domain of cDNA clone 45. The genomic sequence (positions 2501–2518 (Sloof *et al.*, 1987) is aligned with the nucleotide sequence of the corresponding region of the cDNA. Editing sites are numbered from 3′ to 5′. Lower case T's in the cDNA represent inserted U's in the transcript; lower case T's in the genomic DNA represent U's deleted from the transcript. (C) Polyadenylation sites. The nucleotide sequences of four ND8 cDNAs have been aligned with the genomic sequence. Arrows indicate the polyadenylation sites in the respective transcripts. (D) PCR amplification of ND8 cDNA. Reactions have been performed with two primer sets, C84/C82 (lanes 1–4) and C93/C82 (lanes 6–9), with the following templates: mtRNA (lanes 1 and 2, 6 and 7), unedited (lane 3) and edited (lane 8) cDNA. One set of the reactions containing mtRNA as template has been performed without reverse transcriptase (lanes 2 and 7 for the two primer sets, respectively). Minus template controls are shown in lanes 4 and 9. As molecular weight marker, a HinfI digest of pBR 322 DNA, containing fragments of, respectively, 1631, 517, 506, 424, 396, 298, 221, 220, 154, and 75 base pairs in length, has been co-electrophoresed (lane 5).

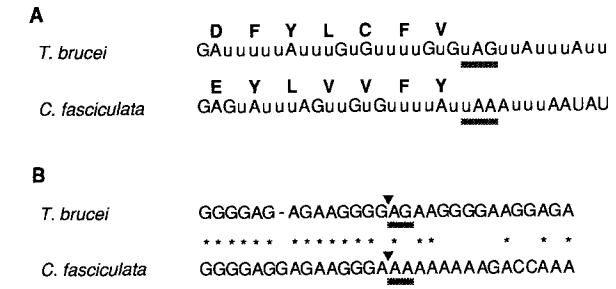

Fig. 4. 3′ End sequences of edited and pre-edited ND8 transcripts of *C. fasciculata* and *T. brucei*. Panel A: Alignment of edited ND8 transcripts of *T. brucei* and *C. fasciculata* with the inferred amino acid sequences. The bar indicates the putative translational termination codon. Panel B: Alignment of GAC sequences (i.e., leaving out the gene encoded U's) of the 3′ terminus of ND8 pre-mRNAs. The bar indicates the position at which in both *T. brucei* and *C. fasciculata* ND8 RNAs a stop codon is created by RNA editing. Asterisks indicate nucleotide identity.

Since the edited segment represents the first (= most 3′) region to be edited, it appears likely that a stretch of 46 nucleotides between the edited sequence and the poly(A) tail would also be unedited in more extensively and fully edited RNAs. Other panedited transcripts including *T. brucei* ND8 RNA, have a 3′ unedited region of a similar size (Feagin *et al.*, 1988a; Koslowsky *et al.*, 1990; Bhat *et al.*, 1990; Souza *et al.*, 1992; Maslov *et al.*, 1992). We took this into account by performing a reverse transcription-PCR analysis with primer pair C84 and C82 (see Fig. 3A). Partially edited ND8 RNAs (if present) should be amplified by this procedure. Only one fragment, however, was produced, the size of which corresponded to that of unedited cDNA (Fig. 3D, lane 1). Surprisingly, the use of primer C93 corresponding to the edited sequence of clone 45 (see Fig. 3A), did not produce any fragment at all in a similar experiment (Fig. 3D, lanes 6 and 7), whereas the primer control experiment with edited cDNA as a template produced a clear band (Fig. 3D, lane 8). These experiments suggest that (more) extensively edited ND8 RNAs are absent from mt RNA of cultured *C. fasciculata* and that the partially edited RNA from which clone 45 was derived is present at such a low concentration that it cannot even be visualized by PCR amplification.

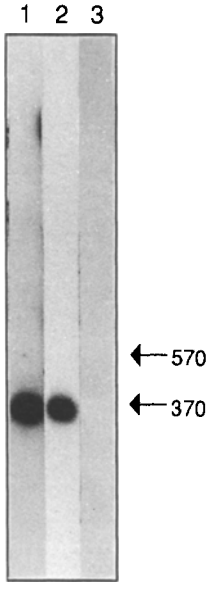

Fig. 5. Northern blot analysis of ND8 transcripts. 1 μg of mt RNA of *C. fasciculata* has been hybridized with a *C. fasciculata* cDNA of the genomic sequence (lane 1), with oligonucleotides C84 (lane 2) and C93 (lane 3). Oligonucleotide C93 hybridized with clone 45 cDNA (not shown). The sequences and coordinates of the oligonucleotides used are given in the Experimental section. The arrows indicate the position of unedited ND8 RNA (370 nucleotides) and the presumed position of edited ND8 RNA (570 nucleotides), respectively.

All further experimental evidence led to the same conclusion: (i) numerous attempts to isolate other edited ND8 cDNAs from the cDNA library using the edited oligonucleotide (C93) as a probe were unsuccessful, (ii) Northern blots of mtRNA probed with unedited cDNA produced one sharp band the size of which corresponds exactly to that of unedited RNA (Fig. 5, lane 1) and not the typical smear of partially edited transcripts that is seen for other panedited RNAs when a similar approach is used (Feagin *et al.*, 1988a; Koslowsky *et al.*, 1990; Bhat *et al.*, 1990; Souza *et al.*, 1992; Maslov *et al.*, 1992). In this case the smear should range in size from 370 nucleotides (unedited ND8 RNA) to about 570 nucleotides, the length of fully edited *T. brucei* ND8 RNA (Souza *et al.*, 1992). Oligonucleotide C84, which as outlined above should hybridize to all ND8 RNAs (unedited, partially edited, and fully edited), produced exactly the same band (Fig. 5, lane 2). Not surprisingly, oligonucleotide C93 gave no signal at all (Fig. 5, lane 3), in agreement with the conclusion that the edited transcript that gave rise to clone 45 is present at a very low concentration. We conclude, therefore, that extensively and completely edited ND8 transcripts are absent in our *C. fasciculata* strain.

An identical approach was followed to identify partially and fully edited transcripts of the ND9 cryptogene in *C. fasciculata*. Also in this case only unedited transcripts were detected (results not shown).

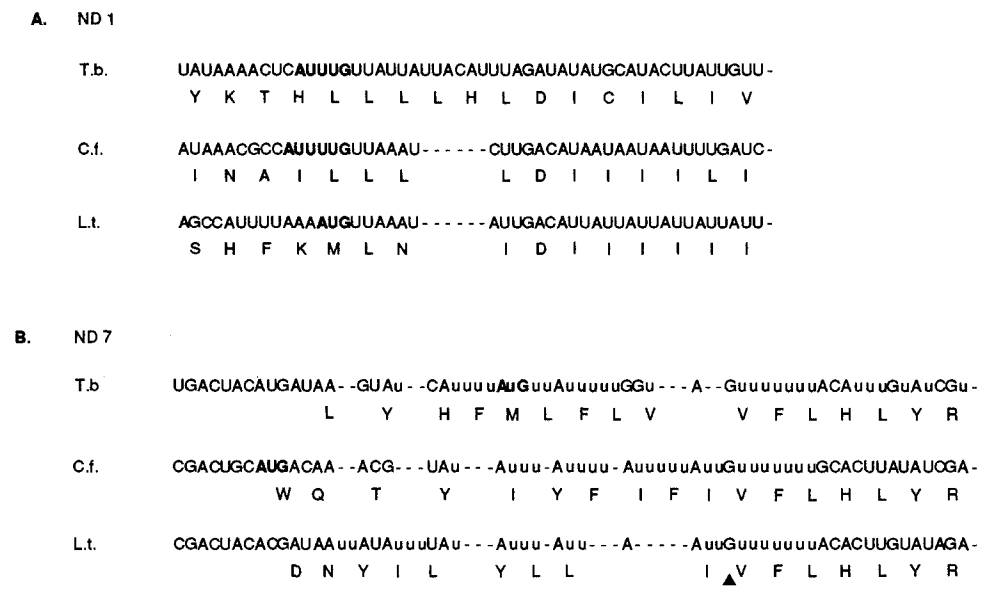

Fig. 6. 5′ End sequences of ND1 and ND7 transcripts in different Trypanosomatid species. Alignment of edited 5′ sequences of ND1 (panel A) and ND7 (panel B) transcripts of *T. brucei* (T.b.), *C. fasciculata* (C.f.), and *L. tarentolae* (L.t.) with the inferred amino acid sequences. Panel A: a potential AUG translation initiation codon in *L. tarentolae* and sites at which AUG codons can be created by RNA editing in the other species are indicated in bold. *T. brucei* ND1 RNA contains 2 extra triplets; gaps are introduced in the sequences of the other organisms to allow optimal alignment. The potential initiator AUG in *L. tarentolae* lines up with UUG triplets in the *T. brucei* and *C. fasciculata* sequences. Panel B: Comparison of the sequences of edited ND7 RNAs in the three species, in which optimal alignment is obtained by introduction of gaps (–). Inserted uridines are given in lower case. Potential AUG translation initiation triplets are in bold; the arrowhead indicates the position upstream of which the editing patterns of the transcripts are divergent.

ND1

The 5′ region of the ND1 transcripts of *C. fasciculata* and *T. brucei* does not contain an AUG triplet that could serve as an initiation codon. However, as outlined in Fig. 6A, they do possess AUUUUG (*C. fasciculata*) and AUUUG (*T. brucei*) sequences, which in theory could be converted into AUG by RNA editing. In *C. fasciculata* this sequence lines up with an AUG codon present at this position in *L. tarentolae* ND1 RNA, and the *T. brucei* motif can be aligned similarly by introducing a six-nucleotide gap in the *L. tarentolae* and *C. fasciculata* sequences (see Fig. 6A). Small gaps must be introduced frequently in the N-terminal region of mitochondrial proteins in order to obtain optimal alignments of homologous sequences in interspecies comparisons (Feagin *et al.*, 1987, 1988b; Shaw *et al.*, 1988; Van der Spek *et al.*, 1990). In fact, the alignment suggested for the ND1 RNAs of the three trypanosome species and the editing proposed in the AUUUUG and AUUUG sequence motifs of *C. fasciculata* and *T. brucei* represents the only way to create an AUG codon in the 5′ region of ND1 RNAs at a similar position in the three species. Extensive experimentation including the analysis of 25 cDNAs and direct sequencing of ND1 RNA has shown that at this location the bulk of the ND1 RNA in cultured *C. fasciculata* is unedited, however (Van der Spek *et al.*, 1990). Therefore, in cultured *C. fasciculata* yet another ND RNA (ND1) is untranslatable.

No information is available on the editing status of ND1 RNA in *T. brucei* (both PF and BF). Since in BF RNA of *T. brucei*, all other ND subunit transcripts are (made) translatable (Stuart, 1993a,b) a translatable ND1 RNA should also be present. In theory, the unedited transcript could be expressed via the use of UUG as an alternative translation initiation codon, as suggested by Sloof *et al.* (1987). Our current prediction would be, however, that edited ND1 versions are present.

ND7

A similar story can be told for ND7 RNA. The pre-mRNA of *C. fasciculata* and *L. tarentolae* is

edited to a limited extent both internally to remove a reading frame shift and in the 5′ region. This 5′ editing does not create an AUG in *L. tarentolae* (Shaw *et al.*, 1989) and shifts an in-frame putative start codon out of phase in *C. fasciculata* (Van der Spek *et al.*, 1988). One cDNA of the *C. fasciculata* collection combined an unedited 5′ region with an edited frameshift region, thus containing a potential AUG initiation codon. It is unlikely, however, that this cDNA represents a functional mRNA because the AUG in question does not align with the putative initiator AUG of the panedited ND7 transcript of BF *T. brucei* (see Fig. 6B). Moreover, the inferred amino acid sequences of the two species are highly divergent in the extreme N-terminal region. Inspection of the ND7 transcript sequences in *L. tarentolae* and *C. fasciculata* indicates that editing cannot create an AUG at a position that would correspond to the start codon of the *T. brucei* sequence (see Fig. 6B). In fact, only at one location can an AUG codon be created by editing at the same position in the three species (the site 5′ of which the inferred amino acid sequences abruptly diverge, indicated by the arrowhead in Fig. 6B). Editing at this position has not been found, however, in any of the cDNAs analyzed in *L. tarentolae* and *C. fasciculata* nor, for that matter, in *T. brucei* (Koslowsky *et al.*, 1990). ND7 therefore appears to represent another example of an ND subunit the production of which is impaired in cultured trypanosomatids.

DISCUSSION

In this study we have listed evidence that transcripts of cryptogenes for ND subunits are not edited into translatable mRNAs in cultured trypanosomatids. The lack of (sufficient) editing in *T. brucei* PF ND RNAs could be explained by the absence of certain key gRNAs, the transcription of which could be down-regulated during this life-cycle stage, or, alternatively, they could be preferentially degraded. In theory, similar events could be invoked to explain the lack of (proper) editing in the *C. fasciculata* and *L. tarentolae* ND RNAs. In such a scenario ND RNAs are not edited because the necessary gRNAs are not present. It has even been proposed for *L. tarentolae* that most gRNA genes required for the editing of the transcripts of cryptogenes other than ATP6 and RPS12 have been lost during the many years that the organisms have been cultivated under laboratory conditions. This would turn these cryptogenes into "pseudo-cryptogenes" (Simpson *et al.*, 1993). Several lines of evidence support this hypothesis: (i) An extensive search resulted in the isolation of only three minicircles encoding an "orphan" gRNA, i.e., a gRNA specifying the editing of an unknown pre-mRNA (Maslov and Simpson, 1992). All other minicircles found encoded a gRNA specifying known edited sequences, some of them showing up numerous times in independently isolated clones. The conclusion that no other gRNA genes are present in the minicircle population of this strain seems therefore a valid one. (ii) Minicircle restriction patterns are considerably more complex in recently isolated strains of *L. tarentolae* than those in laboratory strains (Gomez-Eichelmann *et al.*, 1988), but direct sequence analysis of the minicircles in these strains has not (yet) been performed. Although no recently isolated strains of cultured *C. fasciculata* are available, it has been noted that minicircle restriction patterns change in culture (Hoeijmakers and Borst, 1982). This may suggest that similar losses of minicircle classes take place during the cultivation of this organism too.

Protein products of the mitochondrial ND genes are obviously not required for growth in culture (see below). This would also allow the accumulation of mutations in ND cryptogenes and gRNA genes residing in the maxicircle. For example, the ND7 cryptogene in *L. tarentolae* and *C. fasciculata* could be mutated in such a way that an initiation codon can no longer be created by editing at the correct position. Although the 5′ ND7 gRNA gene has a maxicircle location in these organisms and therefore cannot be lost as easily as a minicircle-encoded gRNA, it can also accumulate mutations. Together with the mutations in the cryptogene, the gRNA mutations could be responsible for the creation of the divergent sequences found in the 5′ area of the edited ND7 RNAs. Consequently, the absence of an initiation codon in *C. fasciculata* and *T. brucei* PF ND1 RNA could simply be the result of mutations in the gene itself and not the result of the absence of a gRNA required for editing.

Whatever the reason for the absence of translatable ND RNAs in cultured trypanosomes, it has some interesting implications for the flow of electrons through the respiratory chain in these organisms. If one assumes that as in most other organisms, the mitochondrially encoded ND subunits of the Trypanosomatidae are part of a large multisubunit

NADH-dehydrogenase (complex 1), it could be envisaged that in culture this complex does not require the presence of the mitochondrially encoded subunits for its activity. Although this cannot be ruled out, it appears not very likely given the deleterious effects even of point mutations in mitochondrial ND genes in mammalian cells (for a review, see Wallace, 1993). It is more likely therefore that in cultured trypanosomes complex 1 is absent. Preliminary results indeed show that O_2 consumption in isolated mitochondrial vesicles of *C. fasciculata* is rotenone insensitive, implying that an active complex 1 is not required (D. Speijer *et al.*, unpublished results). Given the fact that in BF *T. brucei* the mitochondrially encoded ND subunits are expressed under conditions in which cytochromes and classical mitochondrial respiration are absent, they most likely participate in alternative oxidation pathways that do not involve complex 1 also in this life-cycle stage (Souza *et al.*, 1992; Priest and Hajduk, this volume). It is likely, therefore, that in trypanosomes a conventional complex 1 is absent altogether. This would put trypanosomes at the level of other lower eukaryotes such as yeast that do not possess a complex 1-like activity either (see Grivell, 1989).

ACKNOWLEDGMENTS

We are grateful to Drs. D. Speijer and to S. Ploeger and D. Evans for sharing unpublished data, to A. de Haan for expert technical assistance, and to D. Zonneveld for artistic skills in preparing the figures. This work was supported by the Netherlands Foundation for Chemical Research (SON), which is subsidized by the Netherlands Foundation for Scientific Research (NWO).

REFERENCES

Abraham, J. M., Feagin, J. E., and Stuart, K. (1988). *Cell* **55,** 267–272.
Arts, G. J., Van der Spek, H., Speijer, D., Van den Burg. J., Van Steeg, H., Sloof, P., and Benne, R. (1993). *EMBO J.* **12,** 1523–1532.
Bakalara, N., Simpson, A. M., and Simpson, L. (1989). *J. Biol. Chem.* **264,** 18679–18686.
Benne, R. (1993). *Sem. Cell Biol.* **4,** 241–249.
Benne, R., Van den Burg, J., Brakenhoff, J. P. J., Van Boom, J. H., and Tromp, M. C. (1986). *Cell* **46,** 819–826.
Bhat, G. J., Koslowsky, D. J., Feagin, J. E., Smiley, B. L., and Stuart, K. (1990). *Cell* **61,** 885–894.
Blum, B., and Simpson, L. (1990). *Cell* **62,** 391–397.
Blum, B., and Simpson, L. (1992). *Proc. Natl. Acad. Sci. USA* **89,** 11944–11948.
Blum, B., Bakalara, N., and Simpson, L. (1990). *Cell* **60,** 189–198.
Blum, B., Sturm, N. R., Simpson, A. M., and Simpson, L. (1991). *Cell* **65,** 543–550.
Borst, P., and Hoeijmakers, J. H. J. (1979). *Plasmid* **2,** 20–40.
Cech, T. R. (1991). *Cell* **64,** 667–669.
Corell, R. A., Feagin, J. E., Riley, G. R., Strickland, T., Guderian, J. A., Myler, P. J., and Stuart, K. (1993). *Nucleic Acids Res* **21,** 4313–4320.
Decker, C. J., and Sollner-Webb, B. (1990). *Cell* **61,** 1001–1011.
Feagin, J. E., and Stuart, K. (1988). *Mol. Cell. Biol.* **8,** 1259–1265.
Feagin, J. E., Jasmer, D. P., and Stuart, K. (1987). *Cell* **49,** 337–345.
Feagin, J. E., Abraham, J. M., and Stuart, K. (1988a). *Cell* **53,** 413–422.
Feagin, J. E., Shaw, J. M., Simpson, L., and Stuart, K. (1988b). *Proc. Natl. Acad. Sci. USA* **85,** 539–543.
Gomez-Eichelmann, C., Holz, G., Beach, D., Simpson, A. M., and Simpson, L. (1988). *Mol. Biochem. Parasitol.* **27,** 143–157.
Grivell, L. A. (1989). *Eur. J. Biochem.* **182,** 477–493.
Hajduk, S. L., Harris, M. E., and Pollard, V. W. (1993). *FASEB J.* **7,** 54–64.
Harris, M. E., Decker, C., Sollner-Webb, and Hajduk, S. L. (1992). *Mol. Cell. Biol.* **12,** 2591–2598.
Hoeijmakers, J. H. J., and Borst, P. (1982). *Plasmid* **7,** 210–220.
Itakura, K., Rossi, J. J., and Wallace, R. (1984). *Annu. Rev. Biochem* **53,** 323–356.
Kleisen, C. M., Borst, P., and Weijers, P. J. (1975). *Biochim. Biophys. Acta* **390,** 155–167.
Koslowsky, D. J., Bhat, G. J., Perrolaz, A. L., Feagin, J. E., and Stuart, K. (1990). *Cell* **62,** 901–911.
Koslowsky, D. J., Bhat, G. J., Read, L. K., and Stuart, K. (1991). *Cell* **67,** 537–546.
Koslowsky, D. J., Riley, G. R., Feagin, J. E., and Stuart, K. (1992). *Mol. Cell. Biol.* **12,** 2043–2049.
Maslov, D. A., and Simpson, L. (1992). *Cell* **70,** 459–467.
Maslov, D. A., Sturm, N. R., Niner, B. M., Gruszinsky, E. S., Peris, M., and Simpson, L. (1992). *Mol. Cell. Biol.* **12,** 56–67.
Pollard, V. W., Rohrer, S. P., Michelotti, E. F., Hancock, K., and Hajduk, S. L. (1990). *Cell* **63,** 783–790.
Pollard, V. W., Harris, M. E., and Hajduk, S. L. (1992). *EMBO J.* **11,** 4429–4438.
Read, L. K., Correll, R. A., and Stuart, K. (1992a). *Nucleic Acids Res.* **20,** 2341–2347.
Read, L. K., Myler, P. J., and Stuart, K. (1992b). *J. Biol. Chem.* **267,** 1123–1128.
Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). *Molecular Cloning, A laboratory manual,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
Sanger, F., Nicklen, S., and Coulson, A. R. (1977). *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
Shaw, J. M., Feagin, J. E., Stuart, K., and Simpson, L. (1988). *Cell* **52,** 401–411.
Shaw, J. M., Campbell, D., and Simpson, L. (1989). *Proc. Natl. Acad. Sci. USA* **86,** 6220–6224.
Simpson, L. (1987). *Annu. Rev. Microbiol.* **41,** 363–382.
Simpson, L., and Shaw, J. M. (1989). *Cell* **57,** 355–366.
Simpson, A. M., Bakalara, N., and Simpson, L. (1992). *J. Biol. Chem.* **267,** 6782–6788.
Simpson, L., Maslov, D. A., and Blum, B. (1993). In *RNA Editing: The Alteration of Protein Coding Sequences of RNA* (Benne, R., ed.), Ellis Horwood, Chichester, UK, pp. 53–85.
Sloof, P., and Benne, R. (1993). *FEBS Lett.* **325,** 146–151.
Sloof, P., Van den Burg, J., Voogd, A., and Benne, R. (1987). *Nucleic Acids Res.* **15,** 51–65.
Sloof, R., De Haan, A., Eier, W., Van Iersel, M., Boel, E., Van Steeg, H., and Benne, R. (1992). *Mol. Biochem. Parasitol.* **56,** 289–300.
Sollner-Webb, B. (1992). *Nature (London)* **356,** 743–744.

Souza, A. E., Myler, P. J., and Stuart, K. (1992). *Mol. Cell. Biol.* **12,** 2100–2107.
Steinert, M., and Van Assel, S. (1980). *Plasmid* **3,** 7–17.
Stuart, K. (1993a). *Sem. Cell. Biol.* **4,** 251–260.
Stuart, K. (1993b). In *RNA Editing: The Alteration of Protein Coding Sequences of RNA* (Benne, R., ed.), Ellis Horwood, Chichester, UK, pp. 25–52.
Sturm, N. R., and Simpson, L. (1990). *Cell* **61,** 871–878.
Sturm, N. R., and Simpson, L. (1991). *Nucleic Acids Res.* **19,** 6277–6281.
Sturm, N. R., Maslov, D. A., Blum, B., and Simpson, L. (1992). *Cell* **70,** 469–476.
Van der Spek, H., Van den Burg, J., Croiset, A., Van den Broek, M., Sloof, P., and Benne, R. (1988). *EMBO J.* **7,** 2509–2514.
Van der Spek, H., Speijer, D., Arts, G. J., Van den Burg, J., Van Steeg, H., Sloof, P., and Benne, R. (1990). *EMBO J.* **9,** 257–262.
Van der Spek, H., Arts, G. J., Zwaal, R. R., Van den Burg, J., Sloof, P., and Benne, R. (1991). *EMBO J.* **10,** 1217–1224.
Wallace, D. (1993). *Trends Genet.* **9,** 128–133.