The 5'-Terminal Region of the Apocytochrome b Transcript in<br>Crithidia fasciculata Is Successively Edited by Two Guide RNAs in the<br>3' to 5' Direction*

(Received for publication, July 10, 1992)

Hiroyuki Sugisaki‡ and Mituru Takenami
From the Institute for Chemical Research, Kyoto University, Uji, Kyoto 611, Japan

We analyzed the chimeric guide RNA (gRNA)-mRNA molecules in Crithidia fasciculata that are predicted to transiently exist in editing of the 5'-terminal domain of apocytochrome b (Cyb) mRNA, by polymerase chain reaction amplification and DNA sequencing, and obtained evidence suggesting that among the 14 editing sites numbered from 3' to 5', the sequence in the 3' half of the sites (3' block) was specified by one guide RNA species (gRNA-I) and that in the remaining half of the sites (5' block) by another guide RNA species (gRNA-II) and that the direction of editing in each block was 3' to 5'. The predicted transition site of editing by two gRNAs was between the first and second U residues from the 3' end within editing site 7. We found that a stretch of the edited sequence in the 3' block of mRNA could form a stable duplex with a stretch immediately upstream of the guide sequence in gRNA-II. The result leads to a successive editing model that the 3' block of pre-edited mRNA is first edited by gRNA-I, and after completion of editing, the 5' portion of gRNA-II pairs with the edited mRNA for editing of the 5' block.

Several species of mitochondrial mRNA in kinetoplastid protozoans such as Crithidia, Leishmania, and Trypanosoma are extensively edited after transcription (Simpson and Shaw, 1989; Benne, 1990; Feagin, 1991; Harland and Simpson, 1991). The location of editing domains, number of editing sites within a single editing domain, and number of U residues to be added or deleted at each editing site are very specific to individual mRNAs. As to the mechanism of RNA editing, a model that small RNAs, named guide RNA or gRNA, specify the sequence alternation has been proposed, based on the finding that the intergenic regions of maxicircles contain some sequences that can pair with mixed RNA sequences if GU pairs are allowed (Blum et al., 1990). This model was experimentally supported by detection of small RNA species with oligo(U) tails that can hybridize to the corresponding regions of the maxicircles and minicircles (Blum and Simpson, 1990; Pollard et al., 1990; Sturm and Simpson, 1990b; Pollard and Hajduk, 1991; Sturm and Simpson, 1991; van der Spek et al., 1991). Identification of chimeric gRNA-mRNA molecules containing U clusters covalently linked at sites of RNA editing further led to a transesterification model (Cech, 1990; Blum et al., 1991; Koslowsky et al., 1991). Nevertheless, the precise mode of action of the gRNA molecules is yet unknown. According to the computer search data, some of a single editing domain is covered by a few different gRNA species (Blum et al., 1990; Pollard et al., 1990; van der Spek et al., 1991), but its molecular mechanism is also an unsettled question.

The 5'-terminal region of the transcript of the apocytochrome b (Cyb) cryptogene in Crithidia fasciculata contains 14 editing sites, including the one that generates the AUG initiation codon (Feagin et al., 1988a). These sites are tentatively numbered in the 3' to 5' direction (see Fig. 3). By computer analysis, two candidate gRNAs, named gRNA-I and gRNA-II, have been assigned for editing of sites 1-4 (3' or upstream block) and sites 5-14 (5' or downstream block), respectively (van der Spek et al., 1991). The existence of two gRNA species in cells has also been demonstrated (van der Spek et al., 1991). The editing pattern is well conserved between Leishmania tarentolae and C. fasciculata, except that mRNA in L. tarentolae contains an additional editing site at the most 3' end and that the number of U residues at the most 5' editing site differs (Feagin et al., 1988a). The computer assignments of the editing blocks by gRNA-I and gRNA-II are very similar in both strains (Blum et al., 1990; van der Spek et al., 1991), but no direct evidence supporting these predictions has been presented.

One approach to assessing the role of two gRNAs in editing of a single domain is to examine the reaction intermediates predicted by the transesterification model. Detection of a series of expected intermediates may provide supporting evidence for this mechanism, even though it is not experimentally proved yet. We thus analyzed the kinetoplast RNA (kRNA) of C. fasciculata by PCR amplification and DNA sequencing, assuming that both putative gRNA-I and gRNA-II undergo transient covalent interaction with mRNA. As a result, we could identify the chimeric molecules of both gRNA-I and gRNA-II that were covalently linked to partially edited mRNA through U clusters. The sequence data suggested that the transition site of mRNA editing from gRNA-I to gRNA-II was between the first and second U residues from the 3' end within editing site 7, in disagreement with the computer prediction, and that the direction of editing in each block was 3' to 5'. We noted that the 5' portion of gRNA-II could form a stable duplex with a stretch of the mRNA sequence edited by gRNA-I. The result can be inter-
preted that mRNA editing by gRNA-II is initiated only after completion of editing by gRNA-I.

**MATERIALS AND METHODS**

**Preparation of kRNA—**Cells of the C. fasciculata strain Cf-C1 were cultured at 28°C with shaking in the Brain Heart Infusion medium (Difco), supplemented with 10 μg/ml hemin and 100 μg/ml streptomycin sulfate. The kinetoplast was purified from late log-phase cells by the Percoll density gradient centrifugation as described (Birkenmeyer and Ray, 1986). The kRNA fraction was isolated from the kinetoplast by SDS lysis, followed by deproteinization with phenol/chloroform. The ethanol-precipitated pellet was suspended in 10 mM Tris-HCl (pH 8.0) and 1 mM EDTA, from which contaminated kinetoplast DNA networks were removed by centrifugation for 2 hr at 25,000 rpm in the Spinco 50/T rotor, and kRNA was recovered from the supernatant by ethanol precipitation.

**PCR Amplification of Chimeric gRNA-mRNA Regions—**Five μg of kRNA and 100 pmol of the 3′ Cyb mRNA-specific oligonucleotide primer (5′-TCCAACACCATACACTC-3′, nucleotides 3305–3321 in EMBL/GenBank accession number X05063 (Sloof et al., 1987)), in 50 mM Tris-HCl (pH 8.4), 6 mM MgCl₂, 1 mM dithiothreitol, and 1 mM each of dNTP were subjected to denaturation for 3 min at 70°C, followed by annealing for 10 min at 37°C and 10 min at 25°C. One thousand units of Moloney murine leukemia virus H⁻ reverse transcriptase (Bethesda Research Laboratories) was added and the first cDNA strand was synthesized at 42°C for 10 min. The reverse transcriptase was inactivated at 92°C for 2 min, and then 100 pmol of the gRNA-I specific primer (5′-ACCTTGACGTTAGAAGATAA-3′, nucleotides 123–101 in Fig. 9B (van der Spek et al., 1991)) or gRNA-II specific primer (5′-AAGTGGTCTAATAACAAAAAG-3′, nucleotides 166–138 in EMBL/GenBank accession number X05063 (Sloof et al., 1987)) were added together with 6 units of Taq DNA polymerase (Perkin-Elmer Cetus). After PCR amplification (30 cycles of 1.5 min at 94°C; 2 min at 50°C, and 3 min at 72°C, followed by a final 10 min at 65°C) in a thermal circle (Perkin-Elmer Cetus), PCR products were isolated by passage through an agarose gel.

**Cloning and DNA Sequencing—**The PCR products were cloned into the HincII site of pUC118 (Viera and Messing, 1987). Single-stranded DNA was prepared from white colonies obtained, and the sequences of the inserts were determined using the Sequenase Kit (U. S. Biochemical Corp.) (Sanger et al., 1977; Viera and Messing, 1987).

**RESULTS**

**Sequence Analysis of Chimeric gRNA-I-mRNA Molecules—**Assuming that gRNA forms a transient cohesive linkage with partially edited mRNA, the junction sequences were analyzed by PCR amplification which uses a 3′ primer complementary to a downstream region of the pre-edited mRNA and a 5′ primer specific to the 5′ portion of gRNA-I. Since the 5′ end of gRNA-I in C. fasciculata has not been determined, we used the computer-predicted sequence (van der Spek et al., 1990) in the region immediately upstream from the guide sequence (Fig. 1A). The PCR products were cloned in pUC118, and 19 chimeric clones were isolated. When the sequences of individual clones were deduced, 13 different sequence patterns were obtained (Fig. 1A). Each clone contained the gRNA-I sequence, linked to various lengths of oligo(U)ₙ in the 5′ portion and the U stretches were further connected to various editing sites of mRNA in the 3′ portion. Since the presence of oligo(U)ₙ tails in the isolated gRNA molecules has been demonstrated (Blum and Simpson, 1990), the chimeric gRNA-mRNA molecules detected can be interpreted as intermediates in activated states during editing process (Blum et al., 1991).

In all the chimeric gRNA-I-mRNA clones, the gRNA sequence was linked to the U clusters at nucleotide position 39 from the 5′ end of the gRNA primer (Fig. 1A). The junction site can be regarded as the transcriptional termination site of gRNAs, for the oligo(U), tails are assumed to be nonencoded (Blum and Simpson, 1990). The other junctions of the U stretches were distributed between the nucleotides just 3′ of editing sites 1 and 7 of partially edited mRNA, although the linkage site found in the majority of clones was that just 3′ of editing site 7.

**Sequence Analysis of Chimeric gRNA-II-mRNA Molecules—**The 3′ primer used for PCR amplification was the same as that used for analysis of the gRNA-I-mRNA junction. The sequence for the 5′ primer was taken from the 5′-terminal sequence of gRNA-II determined by van der Spek et al. (1991). PCR amplified DNA was cloned, and by analysis of 19 chimeric clones, 17 different sequence patterns were obtained (Fig. 2A). As in gRNA-I-mRNA chimeras, each clone contained the gRNA-II sequences in the 5′ portion which were linked to various editing sites of mRNA through U stretches. The junction of gRNA-oligo(U), in the majority of the gRNA-II-mRNA clones (13 out of 19 clones) was found at nucleotide position 53 or 55 from the 5′ end of the gRNA-II primer. Those in the remaining clones were either truncated or contained unexpected nucleotides in the junction (indicated by dots in Fig. 2A). Fluctuation of the junction sequences observed for gRNA-II-mRNA clones may be due to either ambiguity of transcriptional termination or post-transcriptional modification of gRNA-II. Another possibility to be taken into account would be artifacts in the amplification step. The other junction of oligo(U), in the gRNA-II-mRNA clones were mainly distributed between the nucleotide just 3′ of...
though there is no direct evidence that the chimeras are true chimeric gRNA-mRNA molecules are artificial products cleaved mRNA. However, this is unlikely in view of the precise formation by ligation of the 3' end of gRNA with the 5' end of
Koslowsky products, obtained with a

tion regions which gave different se-

sions (Strum and Simpson, 1990a).

that formed a duplex with pre-edited mRNA at the anchor
gRNA-I to gRNA-I1 is between the first and second U residues
in editing site 7, as in Fig. 1B, and the gRNA-II sequence on
the gRNA-I sequence in the chimeric molecules can fold back on
the remaining mRNA sequence, as in Fig. 2B. According to the
transesterification model, the 3'-terminal U of gRNA,
that formed a duplex with pre-edited mRNA at the anchor
site, first attacks mRNA at the first mismatched base and
produces the chimeric molecule by transesterification. The U
stretch in the gRNA molecule then pairs with the guide A or
G residue of the gRNA itself, and the second transesterifica-
tion takes place at the next mismatched base. As can be seen
in Fig. 1A, the oligo(U), in the majority of clones is connected
with the G residue just 3' of editing site 7, but the A residue
at position 39 at the gRNA-I-oligo(U), junction can pair with
the first U from the 3' end within editing site 7 (Fig. 1B). It
is therefore likely that the transition site of editing from
gRNA-I to gRNA-II is between the first and second U residues
from the 3' end within editing site 7. The mRNA moieties
that were connected to oligo(U), in all the clones have com-
pletely been edited. Thus it is evident that the editing reaction
progressively proceeds in the 3' to 5' direction in both the
eediting blocks. This is essentially consistent with the conclu-
sion deduced from analysis of pre-edited-edited mRNA junc-
tions (Strum and Simpson, 1990a).

**DISCUSSION**

As has already been discussed by others (Blum et al., 1991,
Koslowsky et al., 1991), there is a possibility that the identified
chimeric gRNA-mRNA molecules are artificial products formed by ligation of the 3' end of gRNA with the 5' end of
cleaved mRNA. However, this is unlikely in view of the precise
location of the junction points along the editing sites. Al-
though there is no direct evidence that the chimeras are true
intermediates, our sequence data can be well interpreted on
the basis of the transesterification mechanism.

When the two gRNA sequences were assigned on the
mRNA sequence, we noted that the 12-base-long sequence of
edited mRNA, 5'-uGuaGurruGAAG-3', from editing sites 4-6
can pair with both the guide sequence of gRNA-I and the 5'
moiety of gRNA-II (Fig. 3). The gRNA-II molecule does not
carry any region that can form a stable duplex with pre-edited
mRNA. Although the 5' terminus of gRNA-I has not been
determined yet, the region upstream from the guide sequence
of gRNA-I can pair with the mRNA region just downstream
of the 3' block (Fig. 3). Assuming that the regions of gRNA
molecules which can form stable duplexes with mRNA provide
the anchor sites for mRNA editing (Fig. 3), an editing model
emerged is schematically shown in Fig. 4. The editing reaction
on mRNA of the CYb cryptogene is first initiated by base
pairing-mediated recognition of the pre-edited mRNA se-
quence with the anchor sequence of gRNA-I. Followed by
editing of the 3' block with the guide sequence of gRNA-I,
the anchor sequence of gRNA-II recognizes the edited mRNA
sequence and initiates editing of the 5' block. As the sequences
as well as the editing patterns of C. fasciculata and L. taren-
Toiae are very similar (Feagin et al., 1988a; Blum et al., 1990;
van der Spek et al., 1991), the same mechanism should func-
tion in L. tarentoiae. Actually, a similar base pairing can be
constructed between the edited transcript and the putative
gRNA-II molecule, if a single base gap is inserted. Although
the mechanism involved in switching of binding from gRNA-
I to gRNA-II is not known, the edited mRNA sequence can
form a more stable duplex with the anchor sequence of gRNA-
II than that with the guide sequence of gRNA-I (Fig. 3).

As to the mechanism of RNA editing of a single editing
block, both a consecutive editing (Blum et al., 1990) and
random editing models (Decker and Sollner-Webb, 1990; Kos-
lowsky et al., 1991) have been presented. The former proposed
by Blum et al. (1990) is based on the observation that the 5'
portion of each gRNA candidate carries a sequence hybridiz-
able to a site just 3' of each editing block. This model,
however, does not provide any information about the editing
order of the adjacent blocks by different gRNA candidates.
Furthermore, as argued by van der Spek et al. (1991), the
length of homologous stretches assigned for gRNA-II by Blum

**Fig. 2. A, sequence patterns of chimeric gRNA-II-mRNA molecules. PCR
products, obtained with a 5' gRNA-II-specific primer and a 3' mRNA-specific
primer, are cloned, and those in the junction regions which gave different se-
quencer patterns are indicated as RNAs, aligned with respect to the fully edited
mRNA sequence. Unexpected nucleo-
tides found in the gRNA-oligo(U), junc-
tions and the edited mRNA moiety are
indicated by dots, and gaps required for
sequence alignment are hyphenated. B,
the duplex structure proposed for the
sequences of clone 12. Other annotations
to the figure are identical with those
described in the legend to Fig. 1.
et al. (1990) is too small to support stable duplexes. As demonstrated in this paper, the 5' portion of gRNA-II can form a much more stable duplex with the mRNA sequence specified by gRNA-I. According to our model, editing by gRNA-II is initiated only after completion of the gRNA-I-directed editing reaction.

Sturm and Simpson (1990a) quantitatively analyzed the sequences of unedited-edited junctions of CYb and cytochrome oxidase subunit III (COIII) mRNAs in L. tarentolae by PCR amplification which uses the primer set of the 3'-edited and 5'-pre-edited sequences of mRNA. Although the intermediates analyzed were quite different from ours, they obtained the junction patterns supporting the consecutive editing process from 3' to 5' in consistency with our observation. With Cyb mRNA in L. tarentolae, they identified a strong pausing site of editing in editing site 8, which corresponds to editing site 7 of Cyb mRNA in C. fasciculata. This pausing could be due to the transition of editing process from gRNA-I to gRNA-II. If this is the case, the transition sites in the two strains differ just by one U residue.

In the transcripts of mitochondrial genes for COIII (Pagan et al., 1988b), a subunit of NADH dehydrogenase (Koslowsky et al., 1990), and subunit 6 of ATPase (Bhat et al., 1990) in Trypanosoma brucei, RNA editing is known to occur extensively along the RNA molecules, and each transcript is assumed to be composed of many editing blocks. In fact, three gRNA genes which can guide partially overlapped sequences on the COIII transcript have been identified on the minicircle (Pollard et al., 1990). It is quite likely that the sequential editing mechanism that proposed in this paper by analysis of CYb mRNA is involved in editing of such long editing stretches.

REFERENCES
Benne, R. (1990) Trends Genet. 6, 177–181
Bhat, G. J., Koslowsky, D. J., Pagan, J. E., Smiley, B. L., and Stuart, K. (1990) Cell 61, 885–894
Birkenmeyer, L., and Ray, D. S. (1986) J. Biol. Chem. 261, 2362–2368
Blum, B., and Simpson, L. (1990) Cell 62, 391–397
Blum, B., Bakalara, N., and Simpson, L. (1990) Cell 60, 189–196
Blum, B., Sturm, N. R., Simpson, A. M., and Simpson, L. (1991) Cell 65, 543–550
Cech, T. R. (1991) Cell 64, 697–699
Decker, C. J., and Sollner-Webb, B. (1990) Cell 61, 1001–1011
Feagin, J. E. (1990) J. Biol. Chem. 265, 19573–19576
Feagin, J. E., Shaw, J. M., Simpson, L., and Stuart, K. (1988a) Proc. Natl. Acad. Sci. U.S.A. 85, 539–543
Feagin, J. E., Abraham, J. M., and Stuart, K. (1988b) Cell 53, 413–422
Koslowsky, D. J., Bhat, G. J., Perrollaz, 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
Pollard, V. W., and Hajduk, S. L. (1991) Mol. Cell. Biol. 11, 1668–1675
Pollard, V. W., Rohrer, S. P., Michelotti, E. F., Hancock, K., and Hajduk, S. L. (1990) Cell 65, 783–790
Sanger, F., Nicklen, S., and Coulson, A. R., (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 5463–5467
Simpson, L., and Shaw, J. (1989) Cell 57, 355–366
Sloof, P., van den Burg, J., Voogd, A., and Benne, R. (1987) Nucleic Acids Res. 15, 51–65
Stuart, K. (1991) Trends Biol. Sci. 16, 68–72
Sturm, N. R., and Simpson, L. (1990a) Cell 61, 871–878
Sturm, N. R., and Simpson, L. (1990b) Cell 61, 879–884
Sturm, N. R., and Simpson, L. (1991) Nucleic Acids Res. 19, 6277–6281
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
Vieira, J., and Messing, J. (1987) Methods Enzymol. 153, 3–11