Volume 295, number 1,2,3, 97–101 FEBS 10515 December 1991

RNA editing in ATPase subunit 6 mRNAs in *Oenothera* mitochondria

A new termination codon shortens the reading frame by 35 amino acids

Wolfgang Schuster and Axel Brennicke

Institut für Genbiologische Forschung, Ihnestraße 63, D-1000 Berlin 33, Germany

Received 27 August 1991; revised version received 30 October 1991

The open reading frame encoding ATPase subunit 6 in *Oenothera* mitochondria is edited at 21 positions in all cDNA clones investigated. Only one of these events is silent, all others improve similarity between the homologous polypeptides of other species. The introduction of a new UAA termination codon shortens the polypeptide by 35 amino acids to a carboxy terminus conserved in other species. In one of the cDNA clones, an additional editing event was observed resulting in a premature UAA termination codon in the amino terminal region.

RNA editing; *Oenothera* mitochondria; ATPase subunit 6

1. INTRODUCTION

RNA editing in plant mitochondria changes the information content of the RNA molecules by altering C to U, resulting in polypeptides different from those predicted by the genomic DNA [1–3]. The number of nucleotides altered differs, however, between plant species and also between individual genes within one mitochondrion.

The least edited open reading frame described to date in plant mitochondria codes for the alpha-subunit of the ATPase (*atpA*) in *Oenothera* [4]. Four editing sites have been identified in the 1533 nucleotide open reading frame, two of which are silent and two alter the respective encoded amino acid.

Highly edited transcripts in *Oenothera* include the coding sequences of subunit 3 of the NADH-dehydrogenase (*nad3*) [5] and of subunit II of the cytochrome oxidase (*coxII*) [6]. In other plant species, like wheat, mRNAs of subunit 9 of the ATPase (*atp9*) are extensively edited [7,8], while the *atp9* transcripts in *Oenothera* are edited at only four sites [9].

Most of the RNA editing sites described in plant mitochondrial open reading frames to date alter the encoded amino acid. Significantly less editing events usually occur in silent positions and in non-coding regions. The silent editings are often found edited in only a few of the mRNA molecules, whereas many of the non-silent editings are edited in most transcripts. Editing sites differentially altered in individual cDNA clones have been observed in many plant mitochondrial transcripts, but not in the mRNAs coding for ATPase subunit 9. These mRNAs appear to be all completely edited without any unaltered molecules being detectable [7–9].

The gene encoding another subunit of the ATPase, subunit 6 (*atp6*), has been sequenced in the mitochondrial DNAs of a number of plant species and was found to contain divergent amino and carboxy terminal sequences outside a conserved core [10–16]. Protein processing removes the amino terminal region of the ATP6 protein in yeast [17] and presumably also the plant mitochondrial ATP6 proteins [18]. From further amino acid sequence comparisons and analysis of the corresponding nucleotide sequences we reasoned that an additional termination codon resulting from RNA editing, as described for *atp9* transcripts in several plant species [8,9], could lead to homologous carboxy termini in plant mitochondria.

We therefore analysed cDNA sequences of the *Oenothera atp6* coding region and report RNA editing to create a new translation termination codon in *atp6* mRNAs that excludes the genomic-encoded variant amino terminal extension from the open reading frame. One silent and 20 non-silent editing events were found in all cDNA clones investigated. An additional rare editing event results in a termination codon after amino acid 86. This editing is interpreted as one of the editing mistakes also found at low frequency in other mitochondrial transcripts of *Oenothera*.

2. MATERIALS AND METHODS

Preparation of mitochondrial nucleic acids from tissue culture cells of *Oenothera berteriana* has been detailed previously. Independent cDNA clones were selected from a cDNA library established from random primers as described [19]. PCR amplifications were done as outlined in [9] from the following primers (Fig. 1):

Correspondence address: W. Schuster, Institut für Genbiologische Forschung, Ihnestraße 63, D-1000 Berlin 33, Germany.

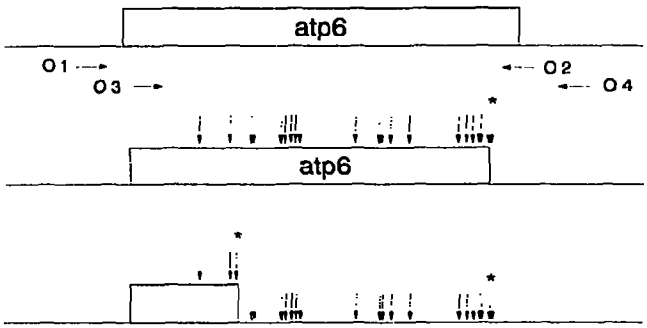

Fig. 1. Schematic representation of the *atp6* gene in *Oenothera* mitochondria and the primers used in PCR amplifications (top line). The sequences of the four oligonucleotides 01, 02, 03 and 04 (horizontal arrows) are given in Section 2. The central part shows the positions of RNA editing sites found in all cDNA clones (vertical arrows). The bottom part represents the truncated open reading frame resulting from the rare editing event observed in the amino terminal region. Stars denote the termination codon introduced by RNA editing events.

01, 5'-CTTGACGAAGTAAAGCTGTCTG-3' (5'-of the open reading frame);
02, 5'-CTTTTGTTCAATTATAAAAAA-3' (at the 3'-end of the open reading frame);
03, 5'-ATAGGTAGTGAAGAAGTAAGT-3' (at the 5'-end of the open reading frame);
04, 5'-AAGAGGGCAATCCCTCGCAC-3' (3'-of the open reading frame).

Varying combinations of primers were employed in these experiments to circumvent spurious PCR products resulting from homologies of 5'-*atp6* sequences with *coxI* and *orfB* [20] and of 3'-*atp6* nucleotides with *nadI* sequences [19].

PCR products were cloned into T-vectors containing a one-nucleotide T-overhang at the *Eco*RV restriction recognition site in Bluescript II (Stratagene) [21]. Nucleotide sequences were determined by the chain termination method with T7 polymerase (Pharmacia) [22].

3. RESULTS

Both independent cDNA clones and PCR-amplified cDNAs were investigated (Fig. 1) in the open reading frame of *atp6*. The entire sequence shown in Fig. 2 was analysed in independent cDNA clones, the sequence between the primers was also determined in cloned PCR products. The nucleotide sequences were identical in all but one of the cDNA clones, but differed consistently in 21 nucleotide identities from the genomic sequence (Figs. 1 and 2).

These differences are in all instances genomic cytidines altered to uridines in the cDNA sequences and are therefore attributed to the RNA editing process in higher plant mitochondria [1–3]. Nineteen of the 21 editing events alter the encoded amino acids. Only one nucleotide at position 810 (Fig. 2) is a silent alteration. The last editing event alters the CAA codon at triplet 261 to a UAA termination codon, resulting in a polypeptide, 35 amino acids shorter than that predicted from the genomic sequence.

The high percentage of non-silent editing sites is reflected in the codon positions of the nucleotides affected. While six editings occur in the first nucleotide and 14 in the second position of the respective codons, only the single silent editing event is a third position alteration.

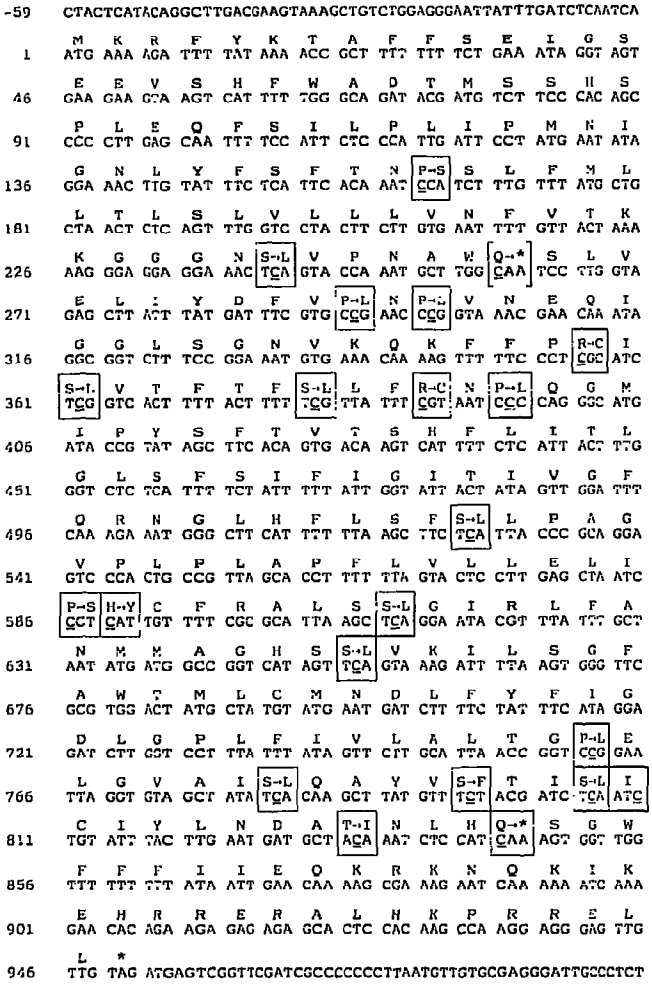

Fig. 2. RNA editing sites in the genomic sequence coding for subunit 6 of the ATPase in *Oenothera* mitochondria. The 21 RNA editing sites observed in all cDNA clones are shown in boxes with the edited nucleotides underlined. The resulting alterations of the codon specificities are indicated; translational stops are marked with asterisks. The editing event at codon 87 observed in only one cDNA clone is shown in brackets. The *atp6* sequence is deposited in the EMBL database under accession number Y00465.

Comparison of the amino acid sequences deduced from the genomic and cDNA sequences of the *Oenothera atp6* gene with the respective polypeptides from other plant and fungal species shows nearly all RNA editings to improve similarity between the species (Fig. 3). Particularly striking is the similarity between the *Oenothera* cDNA sequence and the genomic sequence of radish (*Raphanus*) [15] where only the first editing in *Oenothera* does not yield an amino acid identical to the *Raphanus* genomic-encoded residue.

All editing events occur within the region of homology between the different plant and fungal proteins,

downstream of the codon where the yeast polypeptide has been shown to be processed [17]. The high degree of divergence between the amino terminal extensions in the different plants suggests these likewise to be of no importance for function and probably to be also cleaved off from the mature protein [18]. The first editing occurs far downstream of this putative protein processing site, with all other sites clustered in the open reading frame that is terminated by the last editing site. The 3′-most editing event introduces a new termination codon at the precise termini of the respective open reading frames in *Raphanus* and all fungi as expected from the amino acid comparisons (Figs. 3 and 4). This and many of the other editing events are presumably also required in the monocot species wheat, rice, and maize and in the dicots tobacco, broad bean, and soybean to yield functional ATP6 polypeptides (Fig. 3).

An additional rare editing event was identified in one of the independently derived *atp6* cDNA clones (Fig. 5). This editing alters codon 87 to a UAA termination codon (Fig. 2), disrupting the open reading frame and truncating the encoded polypeptide to a presumably

```
Oenothera cDNA                                                S
                                                                              ↑                74
Oenothera berteriana   MKRFYKTAFFSEIGSEEVSHFWADTMSSHSPLEQFSILPLIPMNIGNLYFSFTNPSLFMLLTLSLVLLLVNFVT
Triticum aestivum     /ETWQGILEDIWAGGGMDNFIQNLPGAYPETPLDQFAIIPIIDLHVGNFYLSFTNEVLYMLLTVVLVVFLFFVVT
Nicotiana tabacum     /EYLLEDLQQNGVGGEAYKDAVDLSKDLVSSPLEQFEIISLIPMKIGNLYFSFTNPSLFMLLTLSLVLLLVYFVT
Oryza sativa          /WQNGIELLDGAEWRNGDIVIPGGGGPVISSPLDQFFIDPLFGLDMGNFYLSFTNESLSMAVTVVLVPSLFGVVT
Zea mays              /DLYDEMERNGEIVNNGSIIIPGGGGPVTESPLDQFGIHPILDLNIGKYYVSFTNLSLSMLLTLGLVLLLVFVVT
Vicia faba            MNLYLQLDLYLYNDNLYLYLYLESGVVPIPSPLEQFEIIPFLPMKIGDLYFSFTNPSLFMLLTLSLVLLLVHFVT
Glycine max  MVEILVTSCWGKQLKNEPKQMKFEWLFLTIAPCAADTINSPLDQFEIIPLIDMKIGDLYFSFTNPSLFMLLTLSLVLLLFHFVT
Raphanus sativus                           MNQIGLVAQSPLDQFEIVPLIPMNIGNFYFSFTNPSLFMLLTLSFFLLLIHFIT
Saccharomyces cerevisiae                  MFNLLNTYITSPLDQFEIRTLFGLQSSFIDLSCLNLTTFSLYTIIVLLVITSLYT
Aspergillus nidulans                        MYQFNFILSPLDQFEIRDLFSLNANVLGNIHLSITNIGLYLSIGLLLTLGYHL
Podospora anserina                   MNTLFNTVNFWRYNSSPLTQFEIKDLISIDTPILGNLHISITNIGFYLTMGAFFLLIINL

Oe cDNA           L     *          L L                  C L     L  C L
                  ↑     ↑          ↑ ↑                  ↑ ↑     ↑  ↑ ↑                        159
Oe. bert.  -KKGGGNSVPNAWQSLVELIYDFVPNPVNEQIGGLSGNVKQKFFPRISVTFTFSLFRNPQGMIPYSFTVTSHFLITLGLSFSIFIG
Tr. aest.  -KKGGGKSVPNAWQSLVELIYDFVLNLVNEQIGGLSGNVKQKFFPRISVTFTFSLFRNPQGMIPFSFTVTSHFLITLALSFSIFIG
Nic. tab.  -KKGGGNSVPNAWQSLVELIYDFVLNPVNEQIGGLSGNVKQKFSPRISVTFTFSLFCNPQGMIPYSFTVTSHFLITLGLSFSIFIG
Oryza sa.  -KKGGGKSVPNAWQSLVELIYDFVLNLVNEQIGG---NVKQKFFPRISVTFTFSLFRNPQGMIPFSFTVTSHFLITLALSFSIFIG
Zea mays   -KKGGGKSVPNAFQSLVELIYDFVPNLVNEQIGGLSGNVKHHFFPCISVTFTFSLFRNPQGMIPFSFTVTSHFLITLALSFSIFIG
Vi. faba   -KKGGGKSVPNAWQSLVELIYDFVPNLVNEQIGGLSGNVKQQFFPCIFVTFTFLLFCNLQGMIPYSFTVTSHFLITLGLSFSIFIG
Gly. max   -KNGGGKLVPNAWQSLVELIYDFVPNLVNEQIAG---NVKQRFFPCISVTFTFSLFCNPQGMIPYSFTVTSHFLITLGLSFSIFIG
Raph. sa.  -KKGGGNLVPNAWQSLVELLYDFVLNLVKEQIGGLSGNVKQMFFPCILVTFLFLLFCNLQGMIPYSFTVTSHFLITLALSFSIFIG
Sacch. c.  LTNNNNKIIGSRWLISQEAIYDTIMNMTKGQIGGK---NWGLYFPMIFTLFMFIFIANLISMIPYSFALSAHLVFIISLSIVIWLG
Asp. nid.  L-AANNKIIPNNWSISQEAIYATVHSIVINQLNPT---KGQLYFPFIYALFIFILVNNLIGMVPYSFASTSHFILTFSMSFTIVLG
Pod. ans.  LSTNYNKLIGNSWSISQESLYATLHSIVVNQINPK---NGQIYFPFIYALFIFILINNLIGMVPYSFASTSHFVLTFALSFTIVLG

Oe cDNA                    L                   SY      L             L
                           ↑                   ↑↑      ↑             ↑                        245
Oe. bert.  ITIVGFQRNGLHFLSFSLPAGVPLPLAPFLVLLELIPHCFRALSSGIRLFANMMAGHSSVKILSGFAWTMLCMNDLFYFIGDLGPL
Tr. aest.  ITIVGFQRHGLHFFSFLLPAGVPLPLAPFLVLLELISYCFRALSLGIRLFANMMAGHSLVKILSGFAWTMLFLNNIFYFIGDLGPL
Nic. tab.  ITIVGFQKNGLHFLSFLLPAGVPLPLAPFLVLLELIPYCFRALSSGIRLFANMMAGHSSVKILSGFAWTMLCMNDLLYFIGDLGPL
Oryza sa.  ITIVGFQRHGLHFFSFLLPAGVPLPLAPFLVLLELISHCFRALSSGIRLFANMMAGHSSVKILSGFAWTMLFLNNIFYFIGDLGPL
Zea mays   ITIVGFQRHGLHFFSFLLPAGVPLPLAPFLVLLELISHCFRALSSGIRLFANMMAGHSSVKILSGFAWTMLFLNNIFYFLGDLGPL
Vi. faba   ITIVGFQRNGLHFLSFLLPAGVPLPLAPFLVLLELISYCFRALSLGIRLFANMMAGHSLVKILSGFAWTMLCMNDLLYFIGDLGPL
Gly. max   ITIVGFQRNGLHFLSFSLPAGVPLPLAPFLVLLELIPHCFRALSSGIRLFANMMAGHSSVKILSGSAWTMLCMNDLFYFIGDLGPL
Raph. sa.  ITIVGFQRHGLHFFSFLLPAGVPLPLAPFLVLLELISYCFRALSLGIRLFANMMAGHSLVKILSGFAWTMLCMNEIFYFIGALGPL
Sacch. c.  NTILGLYKHGWVFFSLFVPAGTPLPLVPLLVIIETLSYFARAISLGLRLGSNILAGHLLMVILAGLTFNFMLINLFTLVFG-FVPL
Asp. nid.  ATFILGLQRHGLKFFSLFVPSGCPLGLLPLLVLIEFISYLSRNVSLGLRLAANILSGHMLLSILSGFTYNIMTSGILFFFLG-LIPL
Pod. ans.  ATILGFQKHGLEFFSLLVPAGCPLGLLPLLVLIEFISYLARNISLGLRLAANILSGHMLLHILAGFTYNIMTSGIIFFFLG-LIPL

Oe cDNA            L      L    F  L        I   *
                   ↑      ↑    ↑  ↑        ↑   ↑                                316
Oe. bert.  FIVLALTGPELGVAISQAYVSTISICIYLNDATNLHQSGWFFFIIEQKRKNQKIKEHRRERALHKPRRELL*
Tr. aest.  FIVLALTGLELGVAISQAHVSTISICIYLNDATNLHQNESFHN*
Nic. tab.  FIVLALTGLELGVAISQAHVSTILICIYLNDAINLHQSASFFIIEQKRV*
Oryza sa.  FIVLALTGLELGVAILQAHVSTISICIYLNDAINLHQNE*
Zea mays   FIVLALTGLELGVAISQAHVSTISICIYLNDATNLHQNESFHNCIKTRSQS*
Vi. faba   FIVLALTGPELGVAISQAHVSTISICIYLNDATNLHQTCLLFIYN*
Gly. max   FIVLALTGPELGVAISQAHVSTISICIYLNDAINLHQTG*
Raph. sa.  FIVLALTGLELGVAILQAYVFTILICIYLNDAINLH*
Sacch. c.  AMILAIMMLEFAIGIIQGYVWAILTASYLKDAVYLH*
Asp. nid.  AFIIAFSGLELAIAFIQAQVFVVLTCSYIKDGLDLH*
Pod. ans.  AFIIAFSGLELGIAFIQAQVFVVLTSGYIKDALDLH*
```

Fig. 3. Amino acid alignment of the ATP6 polypeptides deduced from genomic and cDNA sequences of *Oenothera* and other plant and fungal species. Almost all editings improve similarity of the deduced *Oenothera* ATP6 with other species. All but one of the amino acids specified by the *Oenothera* edited mRNA sequences are already genomically encoded in the *Raphanus* sequence that thus may require no editing. The amino terminal extensions of wheat, tobacco, rice, and maize are not shown in their complete length. Amino acid sequences deduced from genomic DNA sequences are given for *Oenothera berteriana* [16], *Triticum aestivum* [11], *Nicotiana tabacum* [10], *Oryza sativa* [24], *Zea mays* [12], *Vicia faba* [14], *Glycine maximum* [13], *Raphanus sativus* [15], *Saccharomyces cerevisiae* [25], *Aspergillus nidulans* [26] and *Podospora anserina* [27].

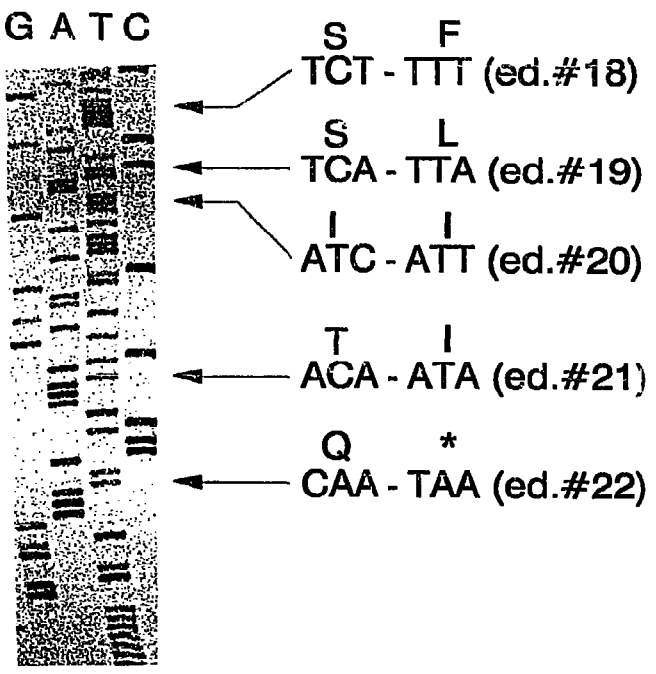

Fig. 4. Several RNA editing events are identified in the cDNA sequence at the 3'-terminus of the *atp6* open reading frame, including the new translation termination codon (editing event #22). The sequence is read from top to bottom in the 5' to 3' direction. Triplets and encoded amino acids are given for the codons altered by RNA editing and the respective edited nucleotide is indicated by an arrow. Numbering of editing sites follows the presentation in Fig. 2.

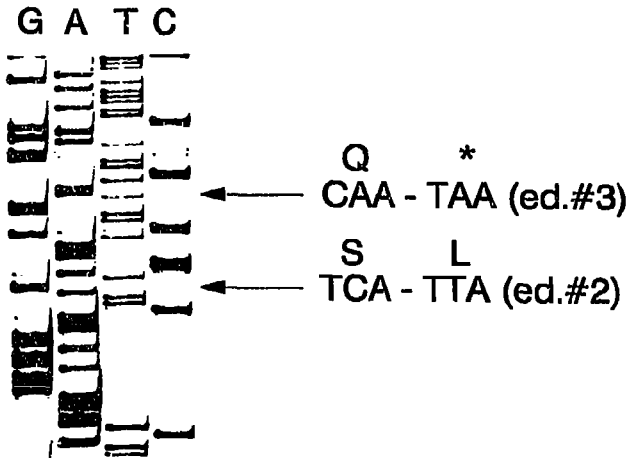

Fig. 5. Documentation of the rare editing mistake resulting in a premature termination codon after amino acid 86 (edit #3; Fig. 3). The sequence reads 5' to 3' from the bottom upwards. Edited codons are given with the altered nucleotide indicated by an arrow. Editing sites are numbered as in Fig. 2.

non-functional size (Figs. 1, 2 and 5). This consequence suggests RNA editing at this position to result from an error of the editing process.

4. DISCUSSION

The investigation of RNA editing in the *atp6* mRNAs reported here shows editing in *Oenothera* mitochondria to introduce a new termination codon. This translational stop removes the non-conserved carboxy terminal amino acids from the open reading frame and aligns the polypeptide terminus with the ATP6 proteins of other species (Fig. 3). An analogous editing event is predicted in mitochondria of other plant species to lead to evolutionarily homologous carboxy termini for the ATP6 polypeptides. The length of the deduced polypeptide is highly conserved in fungi and plants after RNA editing and amino terminal protein processing. This may reflect functional constraints by the assembly of the ATPase complex and its anchoring in the mitochondrial membrane, which requires a well-conserved ATP6 subunit. Subunit 9 is also part of the complex base and is similarly highly conserved in length in plant and fungal species. Intriguingly, RNA editing in mRNAs of this subunit also leads to a new termination codon aligning the ATP9 carboxy termini in different plant species [8,9].

Another termination codon was observed in one of the cDNA clones (Fig. 5) in the amino terminal part of the deduced polypeptide sequence. Since the nucleotide difference resulting in a premature termination codon in the *atp6* open reading frame was only detected in a single cDNA clone, a cloning or reverse transcription artifact cannot be rigorously excluded. However, the nature of the alteration (C to U) and the correct sequences of all cDNA clones at all other positions also suggest that this clone faithfully represents an in vivo sequence variant. Similar infrequent termination codons have also been found in the cDNA clone populations of the *coxI* and *rps3* genes, coding for subunit I of the cytochrome oxidase and ribosomal protein S3 [23]. These events have been interpreted as aberrant editing events, possibly resulting from a misguided editing specificity. Such events may be more frequent in plant mitochondria than previously suspected.

The *atp6* mRNAs possibly represent an extreme divergence of editing patterns between different higher plant species. While the *Oenothera* transcripts have in this investigation been found to be 'normally' edited, the radish *atp6* mRNAs may not be edited at all. Only one of the *Oenothera* editing sites is not encoded in its edited sequence in the radish genome and this site may not be essential for protein function, being divergent in wheat and fungal species. The radish genomic open reading frame also terminates at the translational stop codon introduced by editing in *Oenothera*. Both types of termination codons have now been described as products of RNA editing in *Oenothera* mitochondria; a UGA codon is generated in the *atp9* mRNAs by editing of a CGA codon [9] and a UAA codon was found in this analysis to be generated from a CAA codon. Thus probably any codon position containing a C can be altered by RNA editing in plant mitochondria unless inaccessible by structural constraints.

Acknowledgements: This work was supported by grants from the Deutsche Forschungsgemeinschaft and the Bundesministerium für Forschung und Technologie.

REFERENCES

[1] Hiesel, R., Wissinger, B., Schuster, W. and Brennicke, A. (1989) Science 246, 1632–1634.
[2] Gualberto, J.M., Lamattina, L., Bonnard, G., Weil, J.-H. and Grienenberger, J.-M. (1989) Nature 341, 660–662.
[3] Covello, P.S. and Gray, M.W. (1989) Nature 341, 662–666.
[4] Schuster, W., Ternes, R., Knoop, V., Hiesel, R., Wissinger, B. and Brennicke, A. (1991) Curr. Genet. (in press).
[5] Schuster, W., Wissinger, B., Unseld, M. and Brennicke, A. (1990) EMBO J. 8, 263–269.
[6] Hiesel, R., Wissinger, B. and Brennicke, A. (1990) Curr. Genet. 18, 371–375.
[7] Begu, D., Graves, P.-V., Domec, C., Arselin, G., Litvak, S. and Araya, A. (1990) The Plant Cell 2, 1283–1290.
[8] Nowak, C. and Kück, U. (1990) Nucleic Acids Res. 18, 7164.
[9] Schuster, W. and Brennicke, A. (1990) FEBS Lett. 268, 252–256.
[10] Bland, M.M., Levings III, C.S. and Matzinger, D.F. (1987) Curr. Genet. 12, 475–481.
[11] Bonen, L. and Bird, S. (1988) Gene 73, 47–56.
[12] Dewey, R.E., Levings III, C.S. and Timothy, D.H. (1985) Plant Physiol. 79, 914–919.
[13] Grabau, E., Havlik, M. and Gesteland, R. (1988) Curr. Genet. 13, 83–89.
[14] Macfarlane, J.L., Wahleithner, J.L. and Wolstenholme, D.R. (1990) Curr. Genet. 18, 87–91.
[15] Makaroff, C.A., Apel, I.J. and Palmer, J.D. (1989) J. Biol. Chem. 264, 11706–11713.
[16] Schuster, W. and Brennicke, A. (1987) Nucleic Acids Res. 15, 9092.
[17] Michon, T., Galante, M. and Velours, J. (1988) Eur. J. Biochem. 172, 621–625.
[18] Begu, D., Graves, P.-V., Araya, A. and Litvak, S. (1988) 3rd International Workshop on the Mitochondrial Genome of Higher Plants, Roscoff, 2.
[19] Wissinger, B., Schuster, W. and Brennicke, A. (1991) Cell 65, 473–482.
[20] Hiesel, R., Schobel, W., Schuster, W. and Brennicke, A. (1987) EMBO J. 6, 29–34.
[21] Marchuk, D., Drumm, M., Saulino, A. and Collins, F.S. (1991) Nucleic Acids Res. 19, p. 1154.
[22] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5468.
[23] Schuster, W. and Brennicke, A. (1991) Nucleic Acids Res. 19, in press.
[24] Kadowaki, K.-I., Suzuki, T. and Kazama, S. (1990) Mol. Gen. Genet. 224, 10–16.
[25] Macino, G. and Tzagoloff, A. (1980) Cell 20, 507–517.
[26] Grisi, E., Brown, T.A., Waring, R.B., Scazzocchio, C. and Davis, R.W. (1982) Nucleic Acids Res. 10, 3531–3539.
[27] Cummings, D.J. and Domenico, J.M. (1988) J. Mol. Biol. 204, 815–839.