Volume 258, number 1, 79-83 FEB 07841 November 1989

RNA editing at a splicing site of NADH dehydrogenase subunit IV gene transcript in wheat mitochondria

Lorenzo Lamattina, Jacques-Henry Weil and Jean-Michel Grienenberger

Institut de Biologie Moléculaire des Plantes du CNRS, Université Louis Pasteur, 12 rue du Général Zimmer, 67084 Strasbourg Cedex, France

Received 14 September 1989

Comparison between the sequence of the gene coding for the wheat mitochondrial NADH dehydrogenase subunit IV (nad4) and the cDNA sequence obtained by reverse transcription, using total wheat mtRNA as template, has shown the presence of a uridine residue, not encoded by the genomic sequence, at the exon2-exon3 junction of the spliced transcript. This U creates a non-encoded CUG leucine codon which is essential for maintaining the reading frame, as shown by the conservation of the amino acid sequence of the NAD4 protein in various species. The addition of a U or the specific post-transcriptional conversion of a C to a U could explain this phenomenon.

RNA editing; NADH dehydrogenase subunit IV; Wheat mitochondria; Intron, class II

1. INTRODUCTION

The discovery of the phenomenon of RNA editing upon transcription of protozoan mitochondrial (mt) genes [1-3], was a major breakthrough in the study of the post-transcriptional control mechanisms which regulate gene expression. RNA editing described initially a process resulting in the addition of non-genomically encoded uridine residues [1]. However, it has recently been defined in a broad sense as processes that result in the production of an RNA molecule which differs in nucleotide sequence from the DNA template in the coding regions and where this difference is not the result of the classic splicing mechanisms [4]. Recently, a specific C to U conversion has been described which occurs during the expression of apolipoprotein B producing two alternative transcripts [5,6]. One transcript has a uridine in place of the genomically encoded cytidine and this conversion introduces a stop codon.

Known RNA editing events have been recently reviewed and were identified in protozoan mitochondria as well as in mammalian organisms [4,7]. In this work, we show the presence of a uridine residue not encoded by the genomic DNA, at the exon2-exon3 junction of the transcript of the split nad4 gene of wheat mitochondria.

This is the first report of such RNA editing in the plant kingdom.

Correspondence address: J.-M. Grienenberger, Institut de Biologie Moléculaire des Plantes du CNRS, Université Louis Pasteur, 12 rue du Général Zimmer, 67084 Strasbourg Cedex, France

2. MATERIALS AND METHODS

2.1. *DNA cloning and sequencing*

Subclones of wheat mtDNA were obtained using M13mp19 vectors after transformation of *Escherichia coli* NM522 host strains [8]. Sequencing was performed using the dideoxynucleotide chain termination method [9]. Inserted fragments of the different sizes were obtained by the ordered deletion method [10].

2.2. *cDNA sequencing*

For cDNA sequencing, a dideoxynucleotide chain termination reaction was developed by modification of a previously published method [11]: 60 μg of total wheat mtRNA were hybridized with 4 ng of the specific ^{32}P-labeled oligonucleotide primer in 250 mM KCl and 10 mM Tris-HCl (pH 8.3). The mix (14 μl) was heated at 80°C for 3 min and the oligonucleotide primer and RNA were annealed for 60-90 min at 5°C below the denaturation temperature of the oligonucleotide. The reactions contain: (a) 24 mM Tris-HCl (pH 8.3), 16 mM $MgCl_2$, 8 mM dithiothreitol, 0.8 mM each (dNTP) and 100 μg/ml actinomycin D; (b) 3 μl of the primer/RNA-template solution; (c) 1 μl of 1 mM either ddATP, or ddCTP or ddGTP, or ddTTP; (d) 8-10 units of avian myeloblastosis virus (AMV) reverse transcriptase. The mix was incubated at 50°C for 45 min. The samples were boiled for 3 min and loaded onto sequencing gel.

2.3. *Hybridization assays*

Southern blot analysis was performed by standard methods [8]. For Northern blot analysis, mtRNA was separated (20 μg in each lane) on a 1.3% agarose vertical gel containing 2.2 M formaldehyde [8]. Staining and blotting of RNA to nitrocellulose filters was as described [12]. The blots were washed twice with 5 × SSPE and 0.5% SDS, at room temperature, twice at 65°C for 15 min in the same solution and once in 0.1 × SSPE and 0.1% SDS at 65°C for 15 min. Blots were exposed overnight at −80°C to X-ray film with an intensifying screen.

3. RESULTS AND DISCUSSION

A complete *Sal*I library of wheat mtDNA was screened with a 820 bp *Bgl*II-*Hin*dIII probe which contains

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/89/$3.50

part of the soybean gene coding for subunit IV of NADH dehydrogenase (mitochondrial complex I) [13]. One clone, named C3 by Quetier et al. [14] was found to hybridize with the soybean probe. This clone contains a 23 kb *Sal*I fragment which is unique in the mitochondrial genome since it does not contain any repeated sequence involved in recombination events [14]. Restriction analysis and Southern hybridization showed that a 1.9 kb *Sac*I fragment (Sac1.9) of C3 (fig.1a) contained the sequence homologous to the soybean probe (data not shown). An open reading frame of 148 codons (orf148) was found in the sequenced fragment. The derived amino acid sequence showed 69% similarity with the C-terminal part of several mitochondrial NAD4-proteins (data not shown). In a Northern experiment, using the Sac1.9 fragment as a probe, we showed that the wheat mitochondrial gene is transcribed into a 1.8 kb RNA. In order to look for the 5′ part of the gene, a synthetic oligonucleotide (O_4) deduced from the DNA sequence of orf148 (fig.1a) was used for the synthesis of a labeled cDNA probe by AMV reverse transcriptase using total wheat mtDNA as template [8]. This cDNA probe hybridized to the Sac1.9 fragment and also to a 2.9 kb *Sac*I fragment (Sac2.9) which is situated upstream of Sac1.9 on the C3 fragment (fig.1a), thus indicating that the nad4 gene is interrupted. The Sac2.9 fragment was cloned in M13mp19 in both orientations and sequenced (data not shown).

In order to cross the intron boundaries and to identify the nucleotide sequence corresponding to the 3′-end of the preceding exon (upstream orf148), the cDNA sequence was determined by the chain termination method (fig.2), using total wheat mtRNA as template. Upon complete analysis of the wheat mtnad4 gene structure (which will be published elsewhere), it was found that orf148 corresponds to the third exon (E3) and that Sac2.9 fragment contains the second exon (E2) of this gene. The two exons are separated by a 3405 nucleotide long intron (I2; fig.1a).

When determined, the cDNA sequence revealed, surprisingly, a difference with respect to the corresponding DNA sequence (fig.2). The transcript contains a uridine at the E2/E3 junction, which is part of a non-encoded CUG leucine codon. This experiment has been repeated 4 times, using different wheat mtRNA extracts, with the same result. As this was the only base change found in the 200 nucleotides sequenced, it is not likely to be the result of a reverse transcriptase error.

The DNA sequences were determined on both strands using cloned mitochondrial DNA. To eliminate any possible artifacts due to the cloning steps, we have also determined the sequence of the genomic wheat mtDNA after amplification by the polymerase chain reaction (PCR), using *Taq* polymerase as described [15,16]. Four synthetic oligonucleotides (O_1–O_4) were used for the synthesis of DNA fragments crossing the intron-exon boundaries (fig.1a). These experiments showed that the wheat mtDNA sequences are strictly identical to the cloned Sac2.9 and Sac1.9 fragment sequences (data not shown). Longer expositions did not give different pattern of hybridization fragments.

We have also considered the possibility that homologous sequences carrying a T could exist in other locations in the wheat mt genome. All Southern hybridizations show that each digest gives hybridizing fragments that are unique when they are probed with either Sac2.9 (fig.1b) or Sac1.9 (data not shown) fragments.

Northern blot hybridizations were performed using 5′-end labeled oligonucleotides [8]. When O_4 and O_5 oligonucleotides were used, corresponding respectively to sequences internal to E3 and to E2, only one transcript of 1800 nucleotides was detected, which therefore represents the mature transcript of nad4 (fig.1c, lanes 1 and 2). Two other oligonucleotides, O_2 and O_6, complementary respectively to the 5′ and 3′ regions of the intron, recognized only one RNA molecule of about 3400 nucleotides, corresponding to the known size of the spliced intron (fig.1c, lanes 3 and 4). These experiments suggest that there exists no other exon sequence into I2 and therefore that the comparison which is made between the DNA and RNA sequence is a valid one. The presence of this U residue is necessary to maintain the correct reading frame as shown by a comparison with already known fungal or mammalian NAD4 (fig.3).

This U not encoded in the gene but found in the wheat nad4 transcript is the first demonstration of RNA editing in plant mitochondria. When analysed, the secondary structure of I2 indicates that it is a class II intron [17]. If one looks 39 nucleotides upstream the 3′ end of this intron, one sees that 44 nucleotides can be folded into the conserved secondary structure of loop V of class II introns (not shown). At the 5′ splicing site, I2 contains the sequence GGGCG with a G instead of the U present in the consensus sequence of class II introns [18]. The 3′-end sequence is Y_4AUC instead of the usual Y_nAU. This could be an indication that the last C residue is the first nucleotide of E3. Therefore, as there is no C in the mRNA at this position, the non encoded U found in the mRNA could be derived from this C.

In the heat mt nad4 transcript, this editing takes place at the junction between exons 2 and 3 and therefore could be a consequence of the splicing mechanism. Up to now, there is no report of a nucleotide insertion at the splicing site in class II introns. The guanidine nucleotide which initiates splicing in class I introns has always been found attached to the excised intron and not inserted between the two ligated exons [19]. It is not known whether another splicing mechanism operates specifically in plant mitochondria.

Whether this editing involves a U addition or a C to U conversion, it would be interesting to determine at what level of the mRNA processing this change occurs.

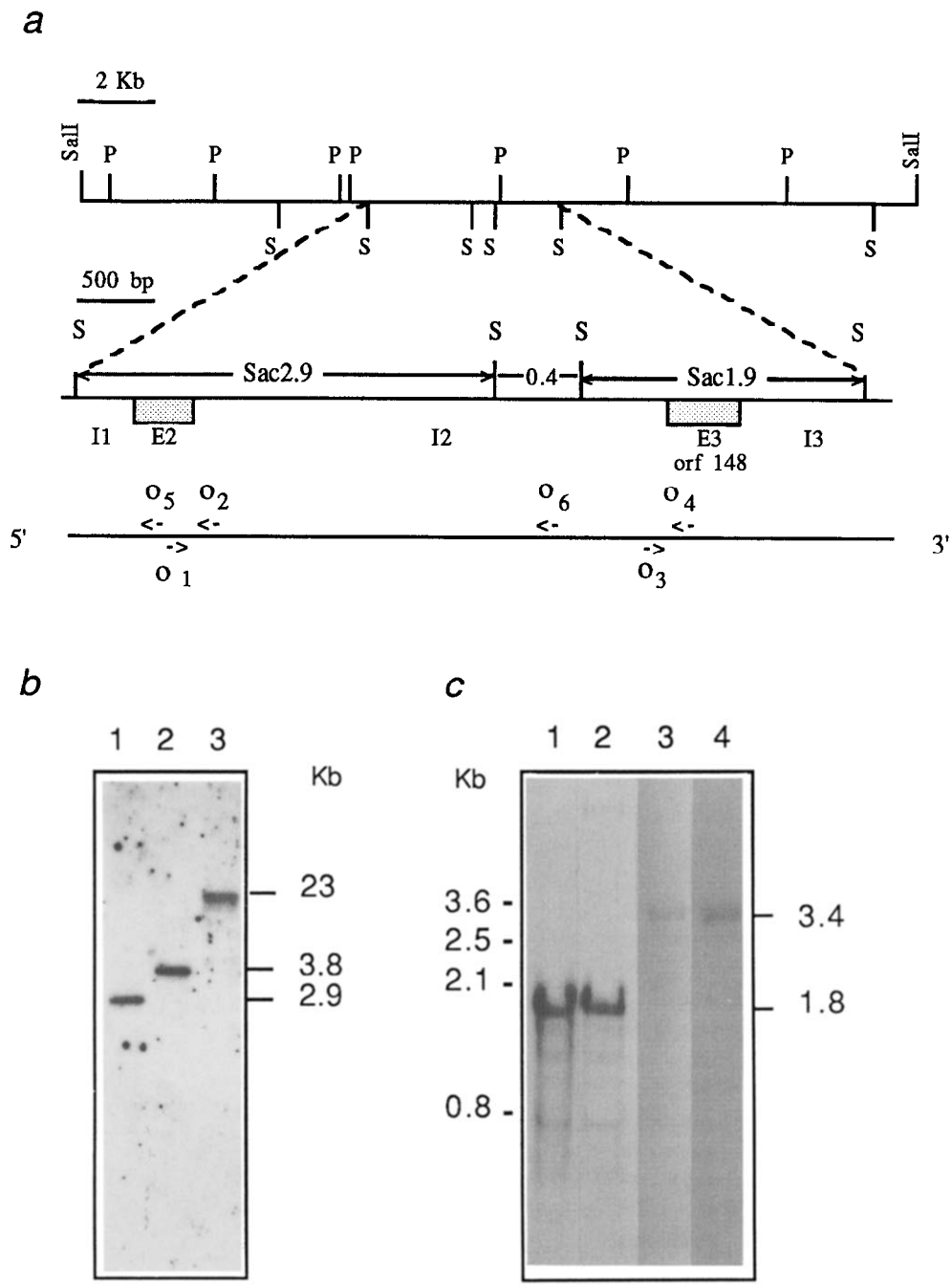

Fig.1. (a) Restriction map of the 23-kb *SalI* fragment (C3) of wheat mitochondrial DNA containing the nad4 gene. I1 (intron 1), E2 (exon 2), I2 (intron 2), E3 (exon 3), I3 (intron 3), O_{1-6}, synthetic oligonucleotides used for the polymerase chain reaction, for DNA and RNA sequencing and for Northern analysis. O_1, 5'CCTTTCATTTATACTCTAAGCGCGATTGC-3'; O_2, 5'-GCGACATGCTAAGTTTTCTCCCCTA-3'; O_3, 5'-TCGATCCGACAGAGCGGAACGACCAG-3'; O_4, 5'-TCCATAATATCTAACAAGTCGAGT-3', O_5, 5'-GGATCCCAGTAAAGTATA-TAGGAAAA-3'; O_6, 5'-CCTTGCTTAGCGTTCCACTTGTG-3'. Relative locations of the primers flanking the edited region are shown with the direction of chain elongation by arrows. P = *PstI*; S = *SacI*. (b) Southern blot analysis of wheat mt genomic DNA restricted with *SacI* (lane 1), *PstI* (lane 2), *SalI* (lane 3) and probed with a second strand synthesis of Sac2.9 fragment cloned in M13. (c) Northern blot analysis of transcripts in total wheat mtRNA probed with O_5 (lane 1), O_4 (lane 2), O_2 (lane 3) and O_6 (lane 4). Numbers at the left indicate the positions of the RNA size markers (BRL).

Unfortunately, it has not been possible to detect the primary transcript corresponding to the nad4 mRNA suggesting that it must be rapidly processed.

The fact that we have been able to detect this editing phenomenon by sequence determination without observing any heterogeneity, indicates that this process modifies all nad4 mRNA molecules and therefore that there probably is a specific mechanism to direct the modification at the precise site of editing. We do not know whether this specificity is encoded in the

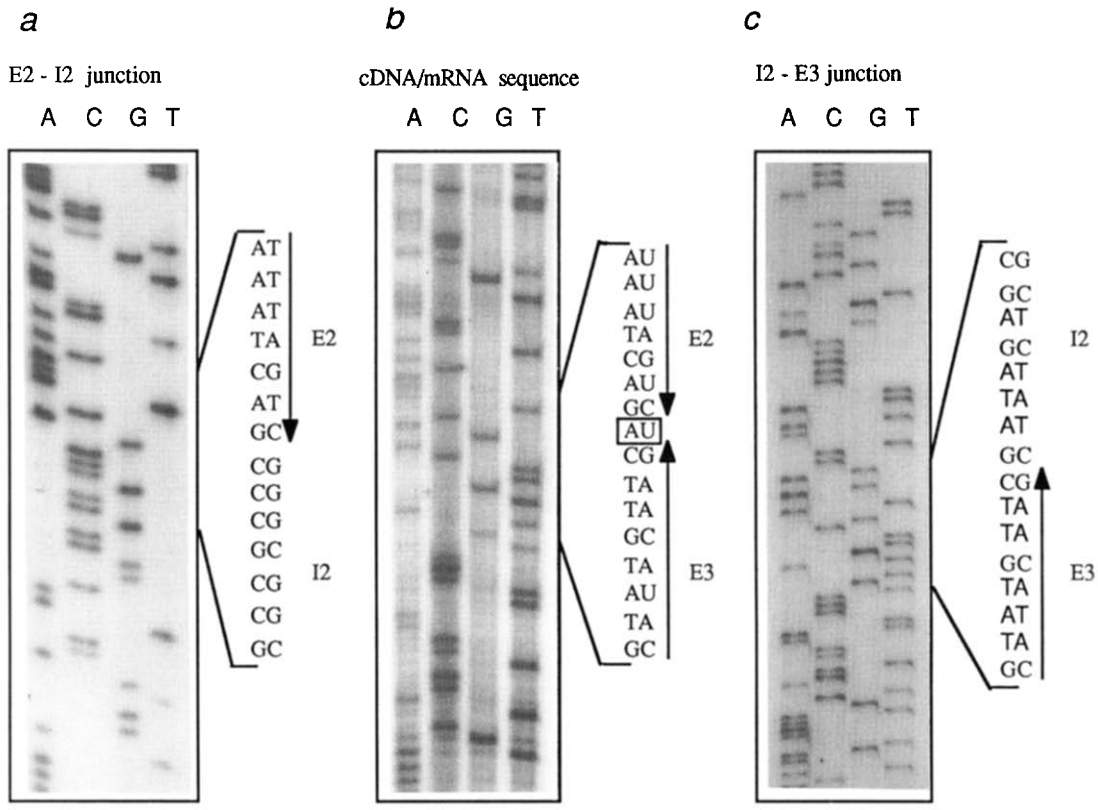

Fig.2. Partial sequences of wheat mitochondrial nad4. (a) DNA sequence of the 3′-end of exon 2 and 5′-end of intron 2. (b) Sequence of nad4 transcript. The box indicates the position of the non-encoded U residue. (c) DNA sequence of the 3′-end of intron 2 and the 5′-end of exon 3. Arrows indicate the extent of exons. (d) The sequence of nad4 transcript is shown next to the DNA sequences corresponding to the exon 2/intron 2 and intron 2/exon 3 boundary regions, aligned above and below, respectively. The box indicates the position of the non-encoded U residue and the nucleotides found at this position in the DNA sequence.

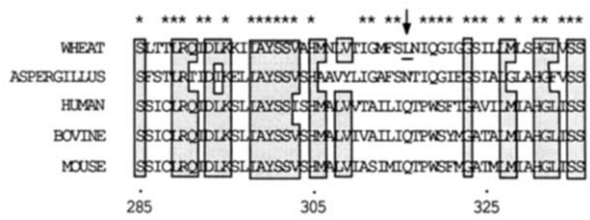

Fig.3. Comparison of part of the deduced amino acid sequence of the wheat mt NAD4 with the amino acid sequence of the corresponding protein from *Aspergillus nidulans* [20], human [21], bovine [22] and mouse [23] mitochondria. The arrow indicates the position of the leucine found at the E2-E3 junction. Stars indicate the homologies with the *Aspergillus nidulans* NAD4 amino acid sequence. Boxes indicate the homologies which exist in this part of the NAD4 protein between wheat and the other organisms.

mitochondrial DNA or due to some structural features of the transcript itself before or after the splicing event.

Acknowledgements: The authors wish to thank Professors F. Quetier and B. Lejeune (Orsay) for providing wheat mtDNA clones, H. Wintz (Cornell University, Ithaca) for the gift of the soybean nad4 probe, J. Gualberto and M. Keller (IBMP, Strasbourg) for many helpful discussions. L.L. is supported by a fellowship from the CONICET (Argentina).

REFERENCES

[1] Benne, R., Van den Brug, J., Brakenhof, J.P.J., Sloof, P., Van Boom, J.H. and Tromp, M.C. (1986) Cell 46, 819–826.
[2] Feagin, J.E., Jasmer, D.P. and Stuart, K. (1987) Cell 49, 337–345.

[3] Shaw, J.M., Feagin, J.E., Stuart, K. and Simpson, L. (1988) Cell 53, 401–411.
[4] Simpson, L. and Shaw, J. (1989) Cell 57, 355–366.
[5] Powell, L.W., Wallis, S.C., Pease, R.J., Edwards, Y.H., Knott, T.J. and Scott, J. (1987) Cell 50, 831–840.
[6] Tennyson, G.E., Sbatos, C.A., Eggerman, T.L. and Brewer, H.B. Jr. (1989) Nucleic Acids Res. 17, 691–698.
[7] Benne, R. (1989) Biochim. Biophys. Acta 1007, 131–139.
[8] Maniatis, T., Fritsch, E. and Sambrook, J. (1982) in: Molecular Cloning: A Laboratory Manual, Cold Spring Harbour Laboratory, Cold Spring Harbor, NY.
[9] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[10] Dale, R.M.K., McClure, B.A. and Houchins, J.P. (1985) Plasmid 13, 31–40.
[11] Hamlyn, P.H., Brownlee, G.G., Cheng, C.C., Gait, M.J. and Milstein, C. (1978) Cell 15, 1067–1075.
[12] Stern, D.B. and Newton, K.J. (1984) Plant Mol. Biol. Rep. 2, 8–15.
[13] Wintz, H., Chen, H.Ch. and Pillay, D.T.N. (1989) Current Genet. 15, 155–160.
[14] Quetier, F., Lejeune, B., Delorme, S., Falconet, D. and Jubier, M.F. (1985) in: Molecular Form and Function of the Plant Genomes (van Vloten-Doting, L., Groot, G.S.P. and Hall, T.C. eds) pp. 413–420, Plenum Press, New York.
[15] Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. and Herlich, H.A. (1988) Science 239, 487–491.
[16] Engelke, D.R., Hoener, P.A. and Collins, F.S. (1988) Proc. Natl. Acad. Sci. USA 85, 544–548.
[17] Michel, F. and Dujon, B. (1983) EMBO J. 2, 33–38.
[18] Cech, T.R. (1986) Cell 44, 207–210.
[19] Sharp, P.A. (1987) Science 235, 766–771.
[20] Brown, T.A., Davies, R.W., Ray, J.A., Waring, R.B. and Scazzocchio, C. (1983) EMBO J. 2, 427–435.
[21] Anderson, S., Bankier, A.T., Barrell, B.G., de Bruijn, M.H.L., Coulson, A.R., Drouin, G., Eperon, I.C., Nierlich, D.P., Roe, B.A., Sanger, F., Schreier, P.H., Smith, A.J.H., Staden, R. and Young, I.G. (1981) Nature 290, 457–465.
[22] Anderson, S., de Bruijn, M.H.L., Coulson, A.R., Eperon, I.C., Sanger, F. and Young, I.G. (1982) J. Mol. Biol. 156, 683–717.
[23] Bibb, M.J., Van Etten, R.A., Wright, C.T., Walberg, W.M. and Clayton, D.A. (1981) Cell 26, 167–180.