Nucleic Acids Research, 1997, Vol. 25, No. 2 **403–409**

RNA editing status of *nad7* intron domains in wheat mitochondria

Catherine Carrillo and Linda Bonen*

Biology Department, University of Ottawa, 30 Marie Curie, Ottawa K1N 6N5, Canada

Received August 13, 1996; Revised and Accepted November 22, 1996

DDBJ/EMBL/GenBank accession no. X75036

ABSTRACT

The most highly conserved structures of group II introns are the helical domains V and VI near the 3′ splice site. Within this region of each of the four introns in the wheat mitochondrial *nad7* gene encoding NADH dehydrogenase subunit 7, there are A-C mispairs. To determine whether C-to-U type RNA editing restores conventional A-U pairing, we sequenced RT–PCR products from partially-spliced *nad7* template RNA and gel-fractionated, excised intron RNA. We examined transcripts from germinating wheat embryos and seedlings because these two stages of development show pronounced differences in steady state levels of *nad7* intronic RNAs. We observed editing at only two of the six predicted sites, and they were located at homologous positions within domain V of the third and fourth introns. A third site was found to be edited within the unmodelled domain VI loop of the fourth intron. Similar patterns of RNA editing were seen in wheat embryos and seedlings. These observations, and the presence of other non-conventional base pairs particularly within domain V of plant mitochondrial introns, indicate weaker helical core structure than in ribozymic group II introns. Moreover, the incompleteness or absence of editing in wheat *nad7* excised intron RNA suggests that, although editing may contribute to splicing efficiency, it is not essential for splicing.

INTRODUCTION

Virtually all plant mitochondrial introns identified to date have been classified as group II and most are located within genes encoding subunits of the NADH dehydrogenase complex (reviewed in 1,2). A subset of these have undergone DNA rearrangements so that mRNA maturation requires splicing *in trans* (reviewed in 3,4). Designation as group II is based on the presence of distinctive secondary–tertiary structural features and such introns are usually depicted as six helical domains radiating from a central wheel (reviewed in 5,6). Over 100 group II introns have been identified in organellar and bacterial genomes and a few, although none in plant mitochondria, have been shown to self-splice *in vitro* (reviewed in 6,7). The catalytic core is comprised of domain V and sequences within domain I, and splicing proceeds through two transesterification steps, the first being initiated by the 2′ hydroxyl group of a bulging adenosine within domain VI. A conformational change between the two steps involves tertiary interactions between sequences in domains VI and II (8). Group II intron excision is mechanistically similar to nuclear pre-mRNA splicing and it has been suggested that these two classes of introns may share a common evolutionary origin.

Domain V is the most diagnostic structure of group II introns and it is essential for splicing *in vitro* and *in vivo*. It is highly conserved in both sequence and length, with a 14 bp helix interrupted by a 2 nucleotide (nt) bulge and closed by a purine tetraloop (reviewed in 5,6). The most conserved bases in this structure, numbered from the 5′-end of the helix, are A2, G3 and C4 which pair with U33, U32 and G31 respectively. The bases A2, G3 and C4 have been shown by mutational analysis to be crucial for splicing (9,10) and the terminal tetraloop, which is usually GAAA, interacts with a motif within domain I (11). The nucleotide sequence of domain V is conservative to such an extent that the identification of group II introns in data bank searches has been possible (12). Domain VI, although less well-conserved and having a shorter helix, almost always possesses a bulging adenosine which is located 7–8 nt from the 3′ splice site and involved in lariat formation.

In addition to intron excision, either *in cis* or *in trans*, the maturation of plant mitochondrial mRNAs also involves C-to-U type RNA editing (reviewed in 2,13,14). It occurs predominantly within protein coding sequences where it typically changes the amino acid specified to one that improves similarity with homologous proteins from other organisms. RNA editing is an early processing event and fully-spliced transcripts are usually fully edited whereas unspliced ones may show partial editing (15,16).

There have been a few RNA editing positions identified within group II intron sequences and it has been suggested that RNA editing may be required to improve base pairing for proper folding needed in splicing (2,3). This model is conceptually appealing because intron editing would then serve an important biological role, as it does in coding regions. Some, but not all, of the RNA editing events observed within plant mitochondrial introns do improve base pairing within the core structure. The conversion of A-C mispairs to A-U pairs has been observed at single sites within the domain I, IV or VI helices of several Oenothera *nad* introns (17–21) and domain I of the potato *rps10*

*To whom correspondence should be addressed. Tel: +1 613 562 5800; Fax: +1 613 562 5486; Email: lbonen@oreo.uottawa.ca

intron (22). Interestingly, editing at a site adjacent to the bulging adenosine in domain VI of the Oenothera *nad1* intron 3 has been found to be necessary for the self-splicing of a chimeric yeast autocatalytic aI5c intron containing the domain VI stem of this Oenothera intron (21). In several other cases, RNA editing weakens helical stability by converting G-C to G-U pairs (17,23), appears not to affect it (24), or editing sites are located outside the defined helical core structure and cannot be assessed (18,24). It should be noted that a number of RNA editing sites have also been observed in the *mat-r* gene which is located within the non-core part of *nad1* intron 4 and they follow editing patterns seen in other plant mitochondrial coding sequences (17,25). No editing sites have previously been reported within the highly conserved domain V.

In our characterization of the wheat mitochondrial *nad7* gene which contains four group II introns, we had predicted that C-to-U type editing would convert A-C mispairs observed within domain V/VI helices to Watson–Crick A-U base pairs (26). In this study, we have used RT–PCR sequencing analysis to examine the RNA editing status of these intronic RNAs in wheat embryos and seedlings.

MATERIALS AND METHODS

Isolation of wheat mitochondrial RNA

Mitochondrial RNA was isolated from 24-h germinating wheat embryos and 6-day etiolated wheat seedlings (*Triticum aestivum* var. Frederick) as previously described (27). To remove residual DNA, RNA preparations were repeatedly precipitated for 2 h with 2 M LiCl and treated with RNase-free DNAse I prior to phenol extraction, phenol–chloroform extraction and ethanol precipitation (28).

cDNA synthesis and PCR amplification

Synthetic oligonucleotide primers were designed from the wheat mitochondrial *nad7* gene sequence (26) and their positions and orientations are shown in Figure 1A. Coordinates are according to the EMBL database accession number X75036.

Primer 1. 5′-CTGGACAAGCTTTAGGGGAA-3′ (552–571)
Primer 2. 5′-ACGCGAATTCGCTTCCGAGG-3′ (1408–1427)
Primer 3. 5′-CTAGGATCCCCACTCATATT-3′ (2705–2724)
Primer 4. 5′-CCACCAGATCTTAAGGAAAG-3′ (4226–4245)
Primer 5. 5′-CCACCACTTCACTTTTGCAC-3′ (4390–4371)
Primer 6. 5′-CTGGTACCTACTGGTACTTC-3′ (4488–4469)
Primer 7. 5′-GCAGTCGACTGAGTTCTGAA-3′ (6155–6174)
Primer 8. 5′-TAGGATCCTGATCGAGCAAG-3′ (6639–6620)

cDNAs were typically generated from ~1 µg wheat mitochondrial RNA template, using AMV reverse transcriptase and 20 ng primer 8 (Fig. 1A). Synthesis was performed according to standard procedures (29) except that RNasin and sodium pyrophosphate were omitted, and the extension was carried out at 50°C. The resulting cDNAs were then amplified using *Taq* DNA polymerase with the appropriate primer pairs. Amplification conditions were 35 cycles of 94°C for 40 s, 52°C for 90 s and 72°C for 90 s.

The RT–PCR products were purified by electrophoresis in TBE–agarose gels and isolation with Gene Clean II (BIO 101) according to supplier's instructions or by spin-recovery with filter

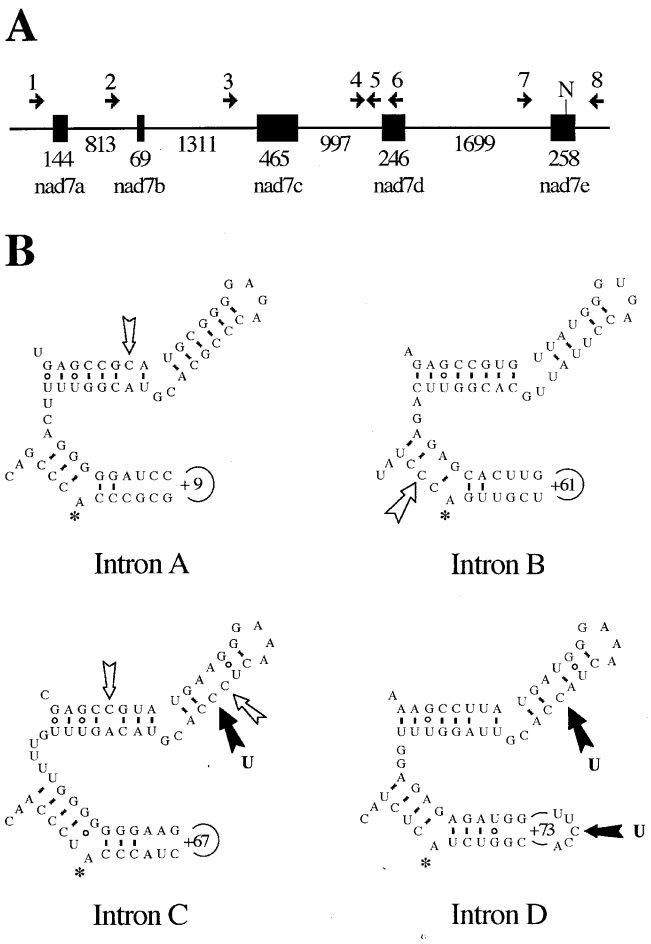

Figure 1. Group II introns in the wheat mitochondrial *nad7* gene. (**A**) Structure of the wheat *nad7* gene with lengths of exons (black bars) and introns (lines) shown in base pairs. Numbered arrows represent orientation and approximate locations of primers used in RT–PCR experiments. The *Nhe*I restriction site (N) shown within exon E is created by RNA editing. (**B**) Secondary structural models of the domain V/VI sequences of the wheat *nad7* introns A–D (updated from ref. 26). The predicted RNA editing sites within the six A-C mispairs are shown by arrows and the RNA editing sites which were experimentally observed are depicted by black arrows. Bulging adenosines within domain VI are shown by asterisks.

tips (30). RT–PCR products were sequenced directly or after being cloned into M13 vectors by standard procedures (28).

Sequencing of RT–PCR products and cloned cDNAs

Samples were sequenced by dideoxynucleotide chain-termination using Sequenase version 2.0 (US Biochemicals). PCR products were sequenced directly using a modification of the Sequenase protocol in which 75 ng primer were annealed with ~500 ng template in the presence of 0.5% Nonidet P-40. Samples were boiled for 3 min, quick-cooled in a –80°C ethanol bath, and then gradually warmed to room temperature. The labelling cocktail was also modified to include four times more labelling mix, ~0.5% NP-40, the supplier's Mn buffer and 1.5 times the suggested amount of dATP. The primers used in sequencing reactions were the same intron primers as in the corresponding PCR reactions, except in the case of the fully-spliced *nad7* mRNA, where an *nad7* exon c primer 5′-GCACAGCAAGCAA-

AGGATTG-3′ (3361–3380, coordinates according to EMBL accession no. X75036) was used.

Isolation and sequence analysis of excised intron RNA

Approximately 10 µg wheat mitochondrial RNA, which had been repeatedly LiCl precipitated, were electrophoresed on a 1.2% agarose gel containing 6.6% formaldehyde in 20 mM MOPS buffer and RNA fractions of interest were isolated using the RNAid kit (BIO 101). The purified RNA was treated with DNase I as above.

cDNA for the intron C region was synthesized using primer 5 and PCR was then performed with primers 4 and 5 as described above except with an annealing temperature of 42°C. To assess the contribution of residual contaminating DNA template in the PCR reaction, we included a set of *nad1a* primers, designated in ref. 31 as A2 (5′-AAGGCTACTCCTAGTAGAAG-3′) and A3 (5′-TCGGGTCGACCAGGTCAGGC-3′), with coordinates 198–179 and 86–105 respectively, according to EMBL accession number X57968. RT–PCR products were gel-purified for direct sequencing as described above. To assess the presence of other *nad7* intron C-containing transcripts, primer 6 was included in RT–PCR reactions with primers 4 and 5, so that any contaminating intron C/*nad7d*-containing transcripts (or DNA) would generate a 263 bp product.

RNA blot analysis

Northern hybridization experiments were performed using standard procedures (28). Approximately 4 µg wheat mitochondrial RNA were electrophoresed on a denaturing formaldehyde gel in MOPS buffer as described above. Hybridization probes were radioactively-labelled by strand-specific primer extension using M13 cloned DNAs, except in the case of introns C and D respectively where the synthetic oligomers 5′-ATGCATGCTTTTGTAGGGTC-3′ and 5′-CGTGTCAGCTTAGTTATC-3′ (EMBL accession no. X75036 coordinates 3494–3475 and 4977–4960 respectively) were radioactively-labelled with T4 polynucleotide kinase and [γ-^{32}P]ATP.

Calculation of predicted thermodynamic stabilities of intron helices

The potential folding energies of domain V helices, constrained to the conventional 14 bp helix with a 2 nt bulge (5), were calculated using the programs available on the 'efn server' web page (http://www.ibc.wustl.edu/~zuker/rna/energy/form1.cgi) which were derived from algorithms described by Zuker *et al.* (32).

RESULTS

Editing status of the domain V/VI regions of wheat *nad7* introns

We had previously established (26) that the wheat mitochondrial *nad7* gene contains four group II introns (Fig. 1A) and has a mature mRNA of ~1.6 kb (Fig. 2F) with 32 C-to-U edits in the coding sequence. When *nad7* intron-specific probes were used in RNA blot analysis (Fig. 2B–E), we observed markedly higher levels of precursor transcripts and excised intron RNAs (Fig. 2, asterisks) in 24-h germinating wheat embryos than in 6-day etiolated seedlings, relative to constant levels of 18S ribosomal

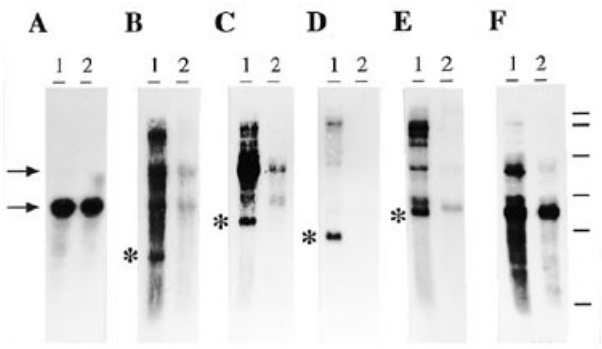

Figure 2. Northern blot analysis of wheat mitochondrial *nad7* transcripts using RNA isolated from 24-h wheat embryos (A–F, lanes 1) and 6-day etiolated seedlings (A–F, lanes 2). Arrows indicate the positions of the 26S and 18S ribosomal RNAs, and bars denote the positions of RNA size markers (from Gibco BRL) of 9.49, 7.46, 4.40, 2.37, 1.35 and 0.24 kb respectively. RNA blots were hybridized with wheat mitochondrial probes for (**A**) 18S rRNA, (**B**–**E**) *nad7* introns A, B, C and D respectively, and (**F**) *nad7* exon E. Asterisks indicate the positions of excised intron RNAs.

RNA (Fig. 2A). In addition, the RNA profiles from these two stages show qualitative differences, most notably in the region of excised intron D RNA species (Fig. 2E asterisk, lanes 1 versus 2). These developmentally-specific differences in the abundance and apparent complexity of *nad7* intronic transcripts may reflect differential RNA stabilities or RNA processing pathways, and raise the possibility of a role for intron editing if such events are important for proper folding for splicing as has been proposed (2,3).

The wheat *nad7* introns have weaker base pairing within the domain V/VI helical regions (Fig. 1B) than do conventional group II introns (cf. Fig. 7A), based on their DNA sequence. There are a total of six A-C mispairs in domain V and VI helices (Fig. 1B, arrows), which could be converted to A-U pairs by RNA editing. Of particular note are two adjacent ones in intron C. It has been estimated, based on NMR spectroscopy, that an A-C mispair within a hairpin is about +2 kcal/mol less stable than an A-U pair (33).

To determine the RNA sequences of these intronic regions, we used a primer mapping downstream of the terminal *nad7* exon (primer 8, Fig. 1A) to generate cDNA from wheat RNA template and then partially-spliced cDNAs were PCR-amplified using specific primer pairs located upstream of the respective domain V sequences of introns A, B and C, as shown in Figure 1A. The observed products, in which all introns except the one of interest had been excised, were 773, 787 and 715 bp in length respectively (Fig. 3A). In the case of the intron D amplification product of 485 bp, an *Nhe*I restriction site within exon E (Fig. 1A) which was created by RNA editing enabled the separation of products derived from edited cDNA molecules as opposed to ones from unedited precursor RNA or residual DNA (Fig. 3B) and this population was examined further. Such experiments were carried out using RNA isolated from 24-h wheat embryos and 6-day seedlings.

For introns A and B, no editing was seen within domain V or VI sequences either by direct RT–PCR sequencing or in any of 10 M13 clones examined for each of these introns. To assess the status of other transcripts containing introns A or B, RT–PCR products were generated using primers 2 and 6 (for intron A) or primers 3 and 6 (for intron B). Direct sequence analysis again

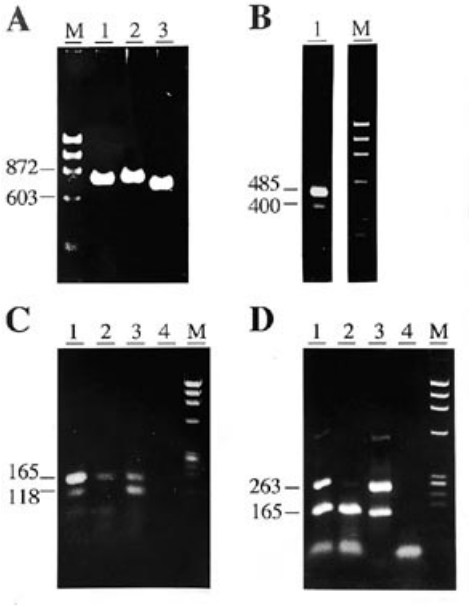

Figure 3. RT–PCR amplification products derived from wheat embryo *nad7* RNA template. (**A**) RT–PCR products from partially-spliced *nad7* template generated using primers 2 and 6 (lane 1), primers 3 and 6 (lane 2) and primers 4 and 8 (lane 3). (**B**) Restriction profile of RT–PCR product generated using primers 7 and 8, after digestion with *Nhe*I, which converts the 485 bp product into 399 bp + 86 bp ones derived from RNA-edited template. (**C**) RT–PCR products generated from excised intron C RNA template using primers 4 and 5 to give a product of 165 bp. The PCR reaction included a second set of primers (A2 and A3 from *nad1a*, ref. 31) which yielded a product of 118 bp, to assess residual DNA contamination. The templates were unfractionated mitochondrial RNA (lane 1), gel-purified excised intron C RNA (lane 2), wheat mitochondrial DNA (lane 3) and a negative control with no DNA template (lane 4). (**D**) RT–PCR products generated from gel-fractionated RNA template using primers 4 and 6 (263 bp product), as well as primers 4 and 5 (165 bp product). The templates were gel-purified RNA from the high MW precursor fraction (lane 1), gel-purified RNA from the excised intron C fraction (lane 2), wheat mitochondrial DNA (lane 3) and a negative control with no template (lane 4). Size markers (M) were ϕX174 DNA restricted with *Hae*III. The positions of primers are shown in Figure 1A.

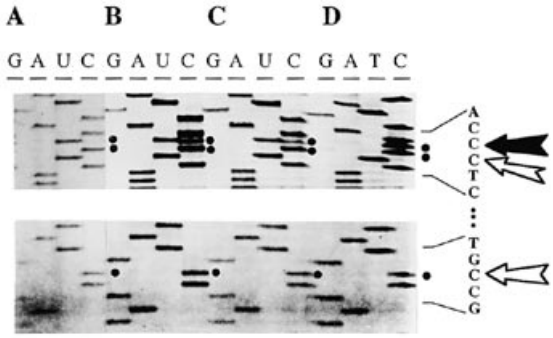

Figure 4. RNA editing status of *nad7* intron C domain V. (A–D) Direct DNA sequencing of RT–PCR products generated from *nad7* intron C wheat embryo template using primer 4 for sequencing. Black arrow indicates the observed RNA editing site whereas open arrows indicate predicted but unedited sites (Fig. 1B). Sequence of RT–PCR products from (**A**) partially-spliced RNA template using primers 4 and 8 (Fig. 3A, lane 3), (**B**) unfractionated RNA using primers 4 and 5 (Fig. 3C, lane 1), (**C**) gel-fractionated excised intron (Fig. 3C, lane 2), (**D**) wheat mitochondrial DNA (Fig. 3C, lane 3).

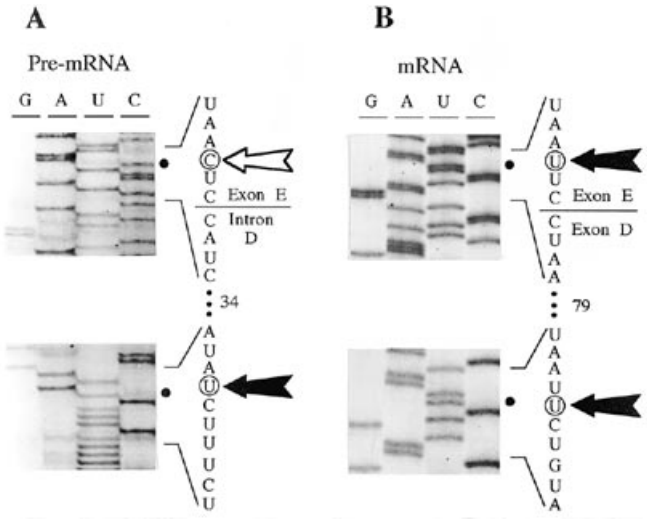

Figure 5. RNA editing status of the *nad7* intron D/exon E junction region. (A,B) Direct DNA sequencing of (**A**) RT–PCR products (primers 7 and 8) restricted with *Nhe*I (Fig. 3B, lane 1) and (**B**) RT–PCR products representing mature *nad7* mRNA (primers 1 and 8). Sequencing primers were primer 7 for (A) and an internal primer mapping to *nad7* exon C for (B). Black arrows indicate observed C-to-U edits, whereas open arrow indicates exon site which is unedited in precursor transcripts. Editing sites within the upstream (A) intron D or (B) exon D are also shown with black arrows.

revealed no intron editing. In the case of intron C, the RNA sequence differed from the DNA sequence at only one position, namely C22 was changed to U22 (Figs 1B and 4A, black arrow), so that an A-C mispair was converted to an A-U pair. At this site, editing was seen in eight out of 10 clones examined for wheat embryo RNA, and six out of 10 clones for seedling RNA. In contrast, at the positions of the other two A-C mismatches (Figs 1B and 4A, open arrows), none of the 20 clones showed editing. In the case of intron D, editing was seen at the position homologous to the edited one in domain V of intron C, and also at a site within the unmodelled loop of domain VI, namely 42 nt upstream of the 3′ splice site (Fig. 5A, black arrow). In wheat embryo RNA, these two intron D sites showed editing in 11/12 and 10/12 clones examined, compared to 8/10 and 7/10 clones respectively when seedling RNA was used as template. Thus, the degree of editing is similar in the two developmental stages, but slightly higher in embryos than in seedlings. With respect to editing within *nad7* coding sequences contained within the RT–PCR products of partially-spliced templates (all of which lacked at least intron D), eight sites were examined in 15–20 cDNA clones and all showed complete or nearly-complete editing (ranging from 70 to 100%). This contrasts with one *nad7e* site, which showed no editing in intron D-containing transcripts (see below).

Because the partially-spliced transcripts showed incomplete editing at the positions of two A-C mispairs and none at the other four predicted sites, we considered the possibility that editing is a rate-limiting step for splicing, and that once editing occurs, the correct structure enables rapid splicing. If so, the population of excised intron RNAs would be expected to be fully edited at those sites. To address this issue, we used gel-fractionated and purified excised intron RNA as template in RT–PCR sequencing experiments. We selected RNA from 24-h embryos because steady state levels are much higher than in seedlings (Fig. 2) and we

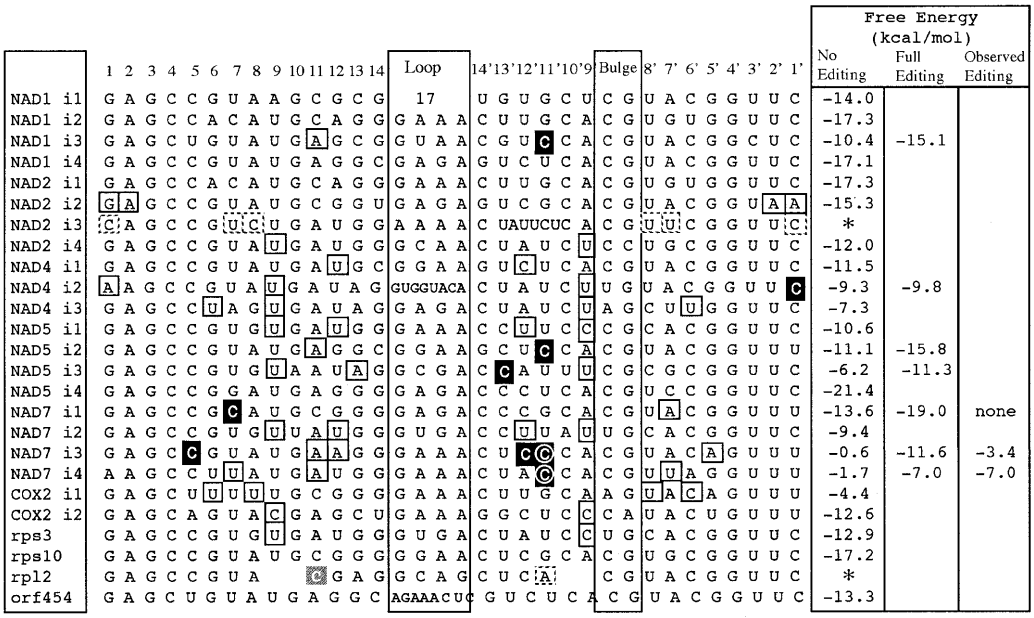

Figure 6. Alignment of domain V sequences from 25 plant mitochondrial group II introns. Sequences are from wheat:*nad1* (EMBL accession nos X57965–X57968), *nad4* (X57164), *nad5* (M74157–M74159) and *nad7* (X75036); Oenothera *nad2* (M81725, M81726), *rps3* (X69140), rpl2 (X80170) and orf454 (X78036); carrot*coxII* (X63625); and potato *rps10* (X74826). The *trans*-spliced introns are: *nad1* i1, i3, i4; *nad2* i2; and *nad5* i2, i3. The 8 bp + 6 bp helices are represented by positions 1–14 and 1′–14′ (see Fig. 7A) and are separated by the dinucleotide bulge. Base pairs other than A-U, G-C or G-U within the helices are blocked and A-C mispairs which could be corrected by RNA editing to A-U are in black blocks. Sites experimentally shown to be edited are also circled. The thermodynamic stabilities of domain V, constrained to the 8 bp + 6 bp helices, were determined using the 'efn server' web page (http://www.ibc.wustl.edu/~zuker/rna/energy/form1.cgi) as described in ref. 32. Values are based on predicted stabilities if there were no editing, the observed editing, or if all sites had been edited (full editing). Asterisks denote the two sequences (*nad2* i3 and *rpl2*) which have length aberrations within the helices and blocks are shown shaded or with broken lines. Free energy values were not calculated for these sequences. If the *cox2* i1 sequence were folded into 6 bp + 8 bp helices (instead of 8 bp + 6 bp), the calculated free energy would be –10.6 kcal/mol.

concentrated on intron C because the excised form is the major intron C RNA species in embryos (Fig. 2D, asterisk) and this intron has three A-C mispairs within domain V, notably two adjacent ones. RT–PCR amplification using intron-specific primers 4 and 5 (Fig. 1A) generated an amplification product of 165 bp (Fig. 3C). To assess the amount of residual contaminating DNA, the PCR amplification step included an additional set of primers which generated a 118 bp product from *nad1a* template DNA (31). The levels of contaminating DNA were seen to be reduced, but not completely eliminated (Fig. 3C, lane 2). To ensure that the excised intron C RNA fraction was not contaminated with other intron C-containing transcripts, cDNAs were synthesized using primers 5 and 6, prior to PCR amplification with primer 4. No products (263 bp in length) derived from intron C/*nad7d*-containing transcripts were observed in the excised intron fraction (Fig. 3D, lane 2), in contrast to their presence in the gel-purified precursor RNA fraction (Fig. 3D, lane 1). Direct sequencing of the *nad7* excised intron C amplification product indicated that only the same single site shows editing (Fig. 4C, black arrow) and that as for the partially-spliced template (Fig. 4A) and unfractionated RNA template (Fig. 4B), it appears to be incomplete in the population of RNA molecules. No editing was detected at the other two positions (Fig. 4C, open arrows).

Editing within the *nad7* terminal exon in precursor and fully-spliced RNA

The RT–PCR product used in the analysis of intron D (see above) contained four exon *nad7e* editing sites (26). One of these, located three nucleotides downstream from the intron–exon border appeared to be unedited as judged by direct sequence analysis (Fig. 5A, open arrow) in contrast to the full editing seen in completely-spliced *nad7* mRNA (Fig. 5B, black arrow) and the nearly-complete editing seen at the intronic site within the loop of domain VI (Fig. 5A, black arrow). Moreover, none of 22 clones examined from embryo or seedling RNAs showed editing at this position. In contrast, the other three exon sites showed >50% editing in these RNAs, which would include early unspliced precursors, but all of which were known to have the *Nhe*I site generated by editing. These observations are consistent with editing being a relatively 'late' RNA processing event at that exon site, and suggest that the context of sequences around the exon junction may be important for editing and/or splicing.

DISCUSSION

We observed C-to-U type RNA editing at only three sites within the domain V/VI regions of the four wheat *nad7* introns. One occurs within the unmodelled domain VI loop in intron D, and the other two convert A-C mismatches to A-U pairs at identical positions within the domain V helix of introns C and D. This is the first report of RNA editing within domain V of any plant mitochondrial intron. Although these two events improve helicity and so may be important for splicing, they do not generate the structure that is invariably present in this catalytic domain of ribozymic group II introns. Moreover, incomplete editing of excised intron RNA indicates that such editing events are not essential for splicing. In addition, four other A-C mispairs within these helices show no editing, so Watson–Crick base pairing at those positions must not be required for proper folding for splicing.

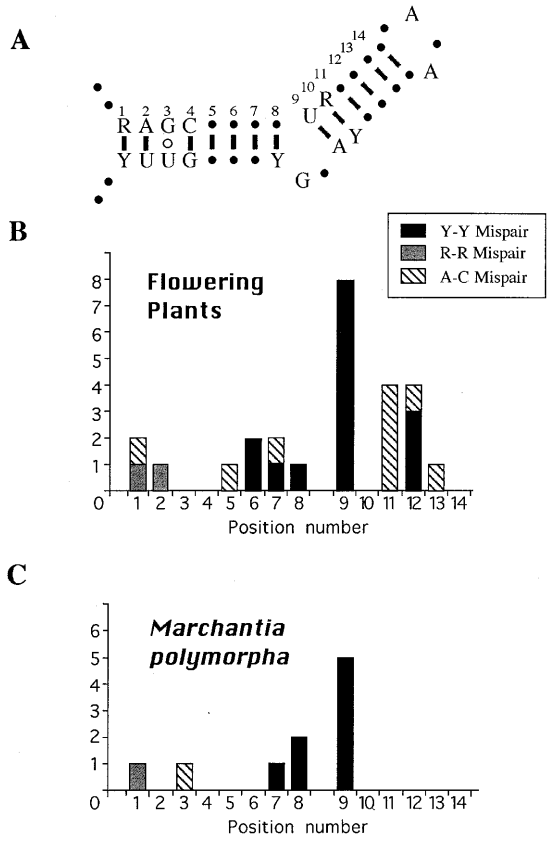

Figure 7. Non-canonical base-pairing within domain V helical positions 1–14 in plant mitochondrial introns. (**A**) Conventional domain V helical structure (according to ref. 5) showing invariant and highly conserved positions (Y, pyrimidine; R, purine). The 8 bp + 6 bp helices are numbered as positions no. 1–14 extending from the central wheel to the loop. (B–C) Location of mispairs within domain V helices of (**B**) 23 flowering plant mitochondrial introns (cf. Fig. 6, with the omission of *nad2* i3 and *rpl2*) and (**C**) 25 group II introns from *Marchantia polymorpha* mitochondria (39). Hatched bars represent A-C mispairs, shaded bars indicate purine–purine mispairs, and black bars show pyrimidine–pyrmidine mispairs.

In the case of intron C, the single editing event shifts the calculated free energy of domain V from –0.6 to –3.4 kcal/mol, which is still markedly less stable than those of ribozymic introns, which have values of –13.0 to –17.9 kcal/mol. Editing at the other two A-C mismatches (Fig. 1B, open arrows) would have resulted in a free energy of –11.6 kcal/mol. Interestingly, this intron which has such weak domain V structure in wheat mitochondria, is absent from the *nad7* gene of tobacco although the other three are present (34). For intron D, the edited helix is considerably more stable (free energy of –7.0 kcal/mol) than the unedited form (–1.7 kcal/mol). The domain V sequences of introns A and B, which showed no editing, have somewhat more stable helices, with free energies of –13.6 and –9.4 kcal/mol respectively.

We observed similar patterns of intron editing in 24-h wheat embryos and 6-day seedlings, with the degree of editing being slightly higher in embryos. Thus, there appears to be no obvious correlation between editing status and the markedly different profiles of *nad7* intron-containing transcripts seen between these two stages of development. Differences in the completeness of RNA editing have been reported for maize *nad3* transcripts, with an increased level seen in 7-day seedlings compared to 3-day seedlings (35). Although it is not known what confers editing site selection in plant mitochondria, it is possible that the editing observed within the domain V helix, designated as base-pair position no. 11 (cf. Fig. 7A) in introns C and D, involves an internal guide RNA. In this regard, editing has also been observed within the acceptor stem of phenylalanine transfer RNA where it improves base pairing and has been shown to be important for tRNA processing *in vitro* (36,37).

In addition to non-canonical A-C mispairs, the *nad7* domain V/VI helices also show several U-U mismatches (Fig. 1B) and such pyrimidine–pyrimidine mispairs are expected to further destabilize helical structure (38). To assess the extent to which plant mitochondrial domain V sequences deviate from the conventional group II structure, we aligned sequences from the 25 different known introns (Fig. 6) and scored sites which did not show the anticipated Watson–Crick (or G-U) base pairing (cf. Fig. 7A). Sequences from several different plants have been used because not all 25 introns have been sequenced in any one plant. When sequence data from other plants were included in the analysis, similar results were obtained because the majority of nucleotide substitutions neither improve nor weaken base pairing (data not shown). Several length variations are seen among the domain V sequences of these introns: three have longer than normal loops, the helix of *rpl2* is shorter by 2 bp, and *nad2* intron 3 has an insert of 3 nt. The latter two were consequently omitted from further analysis. Figure 6 also tabulates the calculated thermodynamic stabilities of the domain V helices when constrained to the classical 14 bp helix. It is evident that only some of the domain V helices with the least favorable free energies are candidates for editing; others have pyrimidine–pyrimidine mispairs.

In Figure 7, the profiles of mispairing within the 14 positions of the domain V helix are shown for the mitochondrial introns of flowering plants and those of the 25 mitochondrial group II introns of the bryophyte, *Marchantia polymorpha* (39), where RNA editing has not been observed nor predicted to be required. It should be noted that only one of these introns, namely *nad2* intron 3, is located at homologous sites in flowering plants and *Marchantia polymorpha.* It is clear that flowering plant mitochondrial introns have more mispairs within domain V helices than those of either *Marchantia polymorpha* or ribozymic group II introns, the latter having no such mispairs. Approximately one-third of the plant domain V mispairs are A-C editing candidates (hatched, Fig. 7B) and the vast majority of the rest are pyrimidine–pyrmidine mismatches. Relaxed base pairing is particularly evident at position no. 9, adjacent to the dinucleotide bulge. On the other hand, positions nos 3 and 4, which are known to be essential for self-splicing *in vitro* (9,10), invariably show conventional pairing. It is notable that at position no. 11, the sole site at which editing was observed in wheat *nad7* domain V sequences, only A-C mispairs which could be corrected by editing, have been observed.

Plant mitochondrial group II introns therefore appear to have a ‘degenerate’ structure, particularly within the domain V helix, when compared to autocatalytic introns. This raises the possibility that they are more dependent on accessory factors, such as proteins and/or small RNAs to achieve splicing. Such proposals have been made for Euglena chloroplast group II introns, which are AT rich and lack certain conventional features (6,40). These issues are also pertinent to the debate over the evolutionary relationship between nuclear spliceosomal introns and group II introns; the latter possess complete splicing information within

the intron RNA, whereas the former require complex small RNA/protein machinery *in trans*. In this regard, it is notable that some plant mitochondrial group II introns are encoded in pieces, which must reassociate at the RNA level for splicing *in trans*. It will be of interest to learn if such constraints play a role in the evolution of group II intron structure in plant mitochondria.

ACKNOWLEDGEMENTS

This work was supported by funding from the Natural Sciences and Engineering Research Council of Canada. The expert technical assistance of K. Williams is acknowledged for data shown in Figure 5B.

REFERENCES

1 Bonen,L. and Brown,G.G. (1993) *Can. J. Bot.* **71**, 645–660.
2 Schuster,W. and Brennicke,A. (1994) *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **45**, 61–78.
3 Wissinger,B., Brennicke,A. and Schuster,W. (1992) *Trends Genet.* **8**, 322–329.
4 Bonen,L. (1993) *FASEB J.* **7**, 40–46.
5 Michel,F., Umesono,K. and Ozeki,H. (1989) *Gene* **82**, 5–30.
6 Michel,F. and Ferat,J.L. (1995) *Annu. Rev. Biochem.* **64**, 435–61.
7 Perlman,P.S. and Podar,M. (1996) *Methods Enzymol.* **264**, 66–86.
8 Chanfreau,G. and Jacquier,A. (1996) *EMBO J.* **15**, 3466–3476.
9 Boulanger,S.C., Belcher,S.M, Schmidt,U., Dib-Hajj,S.D., Schmidt,T. and Perlman,P.S. (1995) *Mol. Cell. Biol.* **15**, 4479–4488.
10 Peebles,C.L., Zhang,M., Perlman,P.S. and Franzen,J.S. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 4422–4426.
11 Costa,M. and Michel,F. (1995) *EMBO J.* **14**, 1276–1285.
12 Knoop,V., Kloska,S. and Brennicke,A. (1994) *J. Mol. Biol.* **242**, 389–396.
13 Gray,M.W. and Covello,P.S. (1993) *FASEB J.* **7**, 64–71.
14 Bonnard,G., Gualberto,J.M., Lamattina,L. and Grienenberger,J.M. (1992) *Crit. Rev. Plant Sci.* **10**, 503–524.
15 Sutton,C.A., Conklin,P.L., Pruitt,K.D. and Hanson,M.R. (1991) *Mol. Cell. Biol.* **11**, 4274–4277.
16 Yang,A.J. and Mulligan,R.M. (1991) *Mol. Cell. Biol.* **11**, 4278–4281.
17 Wissinger,B., Schuster,W. and Brennicke,A. (1991) *Cell* **65**, 473–482.
18 Binder,S., Marchfelder,A., Brennicke,A. and Wissinger,B. (1992) *J. Biol. Chem.* **267**, 7615–7623.
19 Knoop,V. and Brennicke,A. (1993) In Brennicke,A. and Kuck,U. (eds) *Plant Mitochondria.* VCH, Weinheim, pp. 221–232.
20 Lippok,B., Brennicke,A. and Wissinger,B. (1994) *Mol. Gen. Genet.* **243**, 39–46.
21 Borner,G.V., Morl,M., Wissinger,B., Brennicke,A. and Schmeltzer,C. (1995) *Mol. Gen. Genet.* **246**, 739–744.
22 Zanlungo,S., Quinones,V., Moenne,A., Holuigue,L. and Jordana,X. (1995) *Curr. Genet.* **27**, 565–571.
23 Conklin,P.L., Wilson,R.K. and Hanson,M.R. (1991) *Genes Dev.* **5**, 1407–1415.
24 Knoop,V., Schuster,W., Wissinger,B. and Brennicke,A. (1991) *EMBO J.* **10**, 3483–3493.
25 Thomson,M.C., Macfarlane,J.L., Beagley,C.T. and Wolstenholme,D.R. (1994) *Nucleic Acids Res.* **22**, 5745–5752.
26 Bonen,L., Williams,K., Bird,S. and Wood,C. (1994) *Mol. Gen. Genet.* **244**, 81–89.
27 Bonen,L. (1987) *Nucleic Acids Res.* **15**, 10393–10404.
28 Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
29 Krug,M.S. and Berger,S.L. (1996) *Methods Enzymol.* **152**, 316–325.
30 Dean,A.D. and Greenwald,J.E. (1995) *BioTechniques* **18**, 980.
31 Chapdelaine,Y. and Bonen,L. (1991) *Cell* **65**, 465–472.
32 Zuker,M., Jaeger,J.A. and Turner,D.H. (1991) *Nucleic Acids Res.* **19**, 2707–2714.
33 Puglisi,J.D., Wyatt,J.R. and Tinoco,I., Jr. (1990) *Biochem.* **29**, 4215–4226.
34 Pla,M., Mathieu,C., De Paepe,R., Chetrit,P. and Vedel,F. (1995) *Mol. Gen. Genet.* **248**, 79–88.
35 Grosskopf,D. and MulliganR.M. (1996) *Curr. Genet.* **29**, 556–563.
36 Marechal-Drouard,L., Cosset,A., Remacle,C., Ramamonjisoa,D. and Dietrich,A. (1996) *Mol. Cell. Biol.* **16**, 3504–3510.
37 Marchfelder,A., Brennicke,A. and Binder,S. (1996) *J. Biol. Chem.* **271**, 1898–1903.
38 SantaLucia,J., Jr., Kierzek,R. and Turner,D.H. (1991) *Biochemistry* **30**, 8242–8251.
39 Oda,K., Yamato,K., Ohta,E., Nakamura,Y., Takemura,M., Nozato,N., Akashi,K., Kanegae,T., Ogura,Y., Kohchi,T. and Ohyama,K. (1992) *J. Mol. Biol.* **223**, 1–7.
40 Copertino,D.W. and Hallick,R.B. (1993) *Trends Biochem. Sci.* **18**, 467–471.