BIOLOGIA PLANTARUM 52 (2): 347-350, 2008

BRIEF COMMUNICATION

Expression of the *ndhCKJ* operon of barley and editing at the 13th base of the mRNA of the *ndhC* gene

P.H. SERROT, B. SABATER and M. MARTÍN*

Departamento de Biología Vegetal (Fisiología Vegetal), Universidad de Alcalá, Alcalá de Henares, E-28871 Madrid, Spain

Abstract

We have determined a 1778 base sequence which includes the complete *ndhCKJ* operon of barley plastid DNA. This operon contains the *ndhC*, *ndhK* and *ndhJ* genes encoding the polypeptides NDH-C, NDH-K and NDH-J, respectively, of the thylakoid Ndh complex. Poly- and mono-cistronic transcripts were identified, with an increase in the latter under oxidative stress induced by herbicide *Paraquat*. Complete sequencing of transcript cDNAs and of the corresponding regions of five additional monocots revealed that the 13th C (cytosine) base of *ndhC* is edited to U (uracil) converting the CAC codon (encoding histidine, H) to UAC (encoding tyrosine, Y). Dicots having the appropriate TAC codon at the genome sequence do not require editing. The new editing site can not be predicted by comparison with the *Marchantia* sequence (that has a C at the 13th position) because, in contrast to Angiosperms, the amino-terminal sequence in lower plants is highly variable in NDH-C.

Additional key words: herbicide, monocot, *Hordeum vulgare*, oxidative stress, plastid DNA, transcripts.

In many proteomic approaches, genes of data banks which encode specific polypeptides are identified by comparing predicted and determined isoelectric points and sizes of peptides produced by site-specific proteases on purified polypeptides. Many mitochondrial and plastid plant genes are modified by post-transcriptional editing (Maier *et al*. 1996) which changes individual bases in mRNAs thereby modifying the predicted isoelectric points and sizes of peptides and, occasionally, protease cleavage sites. Therefore, the identification of the editing sites in plant organelle genes is necessary for the comparison of determined and predicted peptide features. Most editing in chloroplast changes cytosine (C) to uracil (U) bases at specific positions in transcripts and the screening for candidate editing sites in specific plants is carried out by looking for deviations from conserved amino acids that may be corrected by C→U editing. Comparison of sequences with *Marchantia polymorpha* (which lacks editing machinery) has proved useful (Maier *et al*. 1996). Finally, the editing site is confirmed by determining the cDNA sequence and its comparison with the genomic sequence.

Plastid DNA of most plants contains eleven *ndh* genes encoding polypeptides of the thylakoid Ndh complex which has NADH:plastoquinol oxido-reductase activity and is involved in the poising of the cyclic electron flow (Casano *et al*. 2000), protection against photooxidative stress (Lascano *et al*. 2003) and the induction of leaf senescence (Zapata *et al*. 2005). Almost half of the editing sites identified in plastids are located in *ndh* genes (Sabater *et al*. 2002). The low amount of the Ndh complex in chloroplasts, approximately 1/200 of that of the photosystems (Sazanov *et al*. 1998, Casano *et al*. 2000), makes the conventional identification of its polypeptides difficult and a modern, highly sensitive, proteomic approach requires complete identification of editing sites. With this aim, we have thoroughly investigated the transcripts of the *ndhCKJ* operon of barley (*Hordeum vulgare*) and found a new C→U editing site in the *ndhC* of monocots which is not predicted by comparison with the *Marchantia* genome.

Received 2 June 2006, *accepted* 15 December 2006.
Abbreviations: b - base; PCR - polymerase chain reaction.
Acknowledgements: This work has been supported by Grant BFI2003-04718 of the Spanish DGI (Ministerio de Educación y Ciencia).
* Corresponding author; fax: (+34) 91 8855066, e-mail: mercedes.martin@uah.es

Barley (*Hordeum vulgare* L. cv. Aspen) growth, leaf treatments, the preparation of protein extracts, electrophoresis and immunodetection were carried out as described elsewhere (Casano *et al*. 2001, Lascano *et al*. 2003). DNA and RNA extractions, PCR amplifications, electrophoresis, blotting and hybridization were carried out as described by Del Campo *et al*. (2006).

Both strands of a 1778 base region (accession No. AY243565) of barley plastid DNA were sequenced in an *Applied Biosystem* (Madrid, Spain) automatic sequencer by using PCR (polymerase chain reaction) amplification products. Primers external to the *ndhCKJ* operon (E1: TTCAGTTCCGCCTATTCCAC and E2: GCTCAGTCTATCTGGAATGC) for the primary amplification were derived from the sequence of wheat (accession No. AB042240). Internal primers for secondary amplifications (I1: TAGTGAAGGACCAGA GAAGC; I2: CGCCATGCATAAACTAAACC; I3: AGACCATTCCAAGGCTCCTT; I4: AAGGAGCCTT GGAATGGTCT; I5: CGAATTGGTCAAGACTCTC; I6: CCGCTTGTCTAGGACTTGAT; I7: GTTCCT GTATGAGTACTGCG; I8: GATCACCGAGGAATA GAGAC and I9: CGGCACTTCTCCAAATCCAG) were deduced from our partial sequencing. The sequence obtained for this barley DNA fragment has high homology (around 94 %) with the corresponding regions of rice (accession No. X15901) and maize (accession No. X86563) and lower (but still high, around 83 %) with the corresponding regions sequenced in dicots such as *Arabidopsis* (accession No. AP000423) and tobacco (accession No. Z00044). It spans from 16 bases upstream of the initiation AUG codon of *ndhC* to 89 bases downstream of the stop TGA codon of *ndhJ*. Succesively, the sequence includes the *ndhC* (363 b), *ndhK* (738 b) and *ndhJ* genes (480 b). As in most organisms determined so far, *ndhC* and *ndhK* barley genes overlap by 10 bases at the 3´-end of *ndhC* and 5´-end of *ndhK*. An intergenic 102 base region lies between the 3´-end of *ndhK* and the 5´-end of *ndhJ* of barley which ranges from 97 to 105 bases in other plants.

Northern-blots with probes specific for *ndhC*, *ndhK* and *ndhJ* (Fig. 1*A*) detected several transcripts of the *ndhCKJ* operon. The 1.7 kb transcript was detected with the three probes and must be poly-cistronic including the three gene messages. The 1.1 transcript must include ndhC and ndhK messages and 0.73, 0.48 and 0.36 transcripts must be mono-citronics for *ndhK*, *ndhJ* and *ndhC* genes, respectively. Two bands in the range of 2.0 to 2.6 kb could be attributed to larger poly-cistronic transcripts including neighbour sequences in barley plastid DNA. Significantly, the photo-oxidative treatment by incubation of leaves with *Paraquat* (which increases the production of superoxide anion radical) (Fig. 1, lane *ndhJ*-b) show a strongly increased transcript level, specially those corresponding to the 2.0 to 2.6 kb and the mono-cistronic transcripts but not that of 1.7 poly-cistronic band. A similar behaviour was reported for the multiple transcripts of the *ndhH-D* operon (Del Campo *et al*. 2000, 2002, 2006) indicating a preferential translation of the mono-cistronic transcripts when the Ndh complex is induced by photo-oxidative stress (Casano *et al*. 2000). Effective translation of the *ndhJ* transcript was demonstrated by detection with specific antibody of the encoded protein NDH-J (Fig. 1*B*).

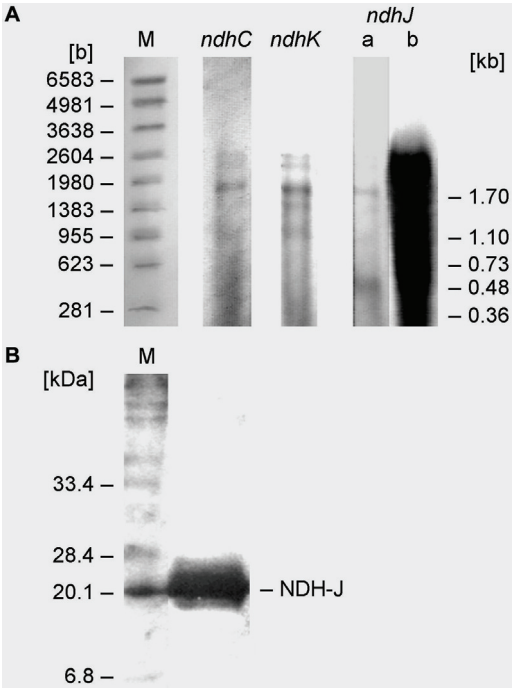

Fig. 1. Expression of the *ndhCKJ* operon. *A* - Northern blots showing transcripts of the *ndhCKJ* operon. Probes specific for the indicated genes were PCR-generated with primers: I1, I2 for *ndhC*; I5, I7 for *ndhK* and I8, I9 for *ndhJ* as described (Del Campo *et al*. 2006). RNA was isolated from young 7-d-old barley leaves and (for *ndhJ b lane*) from 14-d-old barley leaves incubated with 24 h at light with 50 nM *Paraquat*, electrophoresed, blotted and hybridized as described (Del Campo *et al*. 2006, Casano *et al*. 2001). Far-left lane corresponds to size markers. Sizes (in base, *b*) of the main transcripts detected are indicated at right. *B* - Western-blot of a protein extract prepared from 14-days-old barley leaves incubated 24 h at light with 50 nM *Paraquat*. Extract preparation, electrophoresis and immunodetection with NDH-J antibody were performed as described (Lascano *et al*. 2003). The NDH-J antibody was a gift of Dr. W. Kofer (University of Munich, Germany).

To determine all editing sites of barley *ndhC*, *ndhK* and *ndhJ* genes, regions accounting for the entire transcripts were amplified and sequenced from total cDNA. Only one difference between cDNA and genomic sequences was detected: the first base of the 5th codon of the *ndhC* transcript which is C in the genomic sequence and U (determined thymine, T) in the mRNA. Therefore, there is a post-transcriptional C→U editing determining that the fifth codon be translated to tyrosine (Y) instead of histidine (H). We compared it with the genomic sequences described so far around the fifth codon of *ndhC* and, in some cases, determined new genomic and/or cDNA sequences.

Fig. 2*a* shows the barley plastid ndhCKJ operon deduced from sequence determination and the alignment of the first 30 bases of barley *ndhC* with those compiled in NCBI (http://www.ncbi.nlm.nih.gov/) (*Oryza sativa*, *Triticum aestivum*, *Zea mays*, *Spinacia oleracea*, *Atropa belladona*, *Nicotiana tabacum*, *Arabidopsis thaliana*, *Oenothera elata*, *Lupinus luteus*, *Lotus japonicus*, *Picea abies* pseudogene, *Pinus thunbergii* pseudogene, *Psilotum nudum*, *Marchantia polymorpha*, *Anthoceros formosae*, *Chaetosphaeridium globosum*, *Mesostigma viride* and *Nephroselmis olivacea*) and those of *Allium cepa* and *Allium porrum*, also determined in this work.

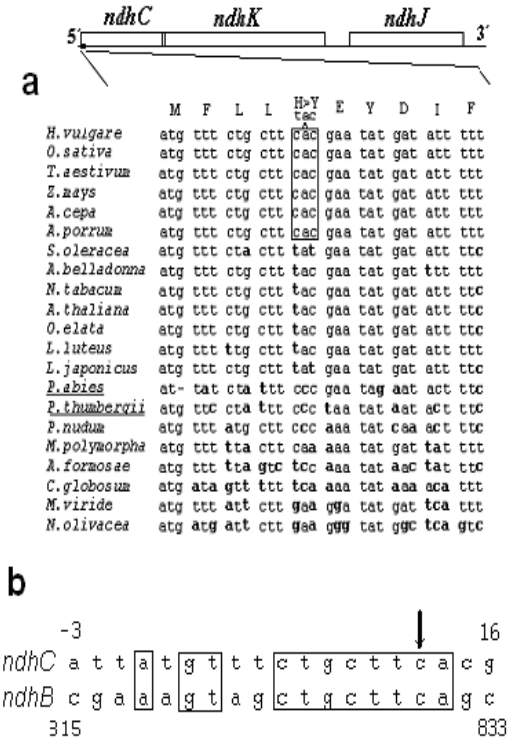

Fig. 2. *a* - Structure of the ndhCKJ operon and alignment of the first 30 bases of plastid *ndhC* genes of different plants. Bases different from those of barley are *in bold*. Amino acids encoded in barley (D - aspartic acid, E - glutamic acid, F - phenylalanine, H - histidine, I - isoleucine, L - leucine, M - methionine, Y - tyrosine) and the monocot editing transition are indicated above base sequences. Note that *P. abies* and *P. thunbergii* have pseudogenes, the sixth codon of the latter being the stop taa. *b* - Alignments of the base sequences surrounding *ndhB*-7 and *ndhC* editing sites (*marked with arrow*). Conserved sequences are *grey boxed*.

All monocots examined (Fig. 2*a*) have a genomic CAC (encoding H) as fifth codon, while all dicots have TAT or TAC (encoding Y). The comparison with *Marchantia* (CAA at the fifth codon) and other lower plants suggests that this position is not conserved in this group. In the place of a functional *ndhC* gene, the Gymnosperms *P. abies* and *P. thunbergii* contain the corresponding pseudogene which, like the *Pteridophyta P. nudum*, have a CCC as the fifth codon. It must be noted that hypothetical correction of the CAA codon of *Marchantia* would produce the TAA stop codon.

We determined cDNA sequences around the fifth codon of the six monocots (Fig. 2*a*) and found that all were edited to TAC, the genomic codon of dicots which encodes the phenolic Y in the place of the basic H. Therefore, the Y at the fifth position of the NDH-C polypeptide sequence seems conserved in Angiosperms and C→U editing restores the codon in monocots. In fact, the first 10 amino acids of NDH-C are conserved in Angiosperms, which contrasts with the high variability in lower plants (Fig. 2*a*). Apparently, structural constraints in the amino-terminal domain of NDH-C appeared in the Angiosperm evolutionary branch that lead to the selection of Y at fifth position.

The lack of functionality of *ndhC* gene in the two Gymnosperms *P. abies* and *P. thunbergii* explains their high sequence difference with other plants (Fig. 2*a*) but, significantly, their fifth pseudocodon is CCC. Furthermore, the 13th base is also C in the Pteridophyta *P. nudum* and the Bryophyta *M. polymorpha* genomes suggesting that the 13th C is an old feature which was modified at the genome level by mutation to T in most lines leading to dicots in order to fulfil new structural requirements at the amino-terminal region of NDH-C. The monocot branch would emerge from dicots conserving a C at the genome level but correcting it by editing to U at post-transcriptional level.

Cavallier-Smith (1997) suggested that editing was a recourse to rescue genes whose functionality had been transitorily lost and had accumulated deleterious mutations. *ndh* genes probably became dispensable in many plants during the Carboniferous period (Sabater *et al*. 2002), when the high CO_2 concentration reduced the severity of photooxidative stress, and consequently accumulated deleterious mutations. *P. abies* and *P. thunbergii* could be remnants of the irreversible loss of functional *ndh* genes. When, at the end of the Carbon, the concentration of atmospheric CO_2 dropped, the Ndh complex again became useful to some plants (such as those leading to Angiosperms) that could rescue the *ndh* genes by post-transcriptional editing of T to C mutations in critical positions. Only later, genomic C to T mutations decreased the number of sites requiring editing. Our results suggest that monocots derived from certain dicots still containing C at the position 13 in *ndhC*. The determination of the 5′end sequence and editing status of the 13th position of *ndhC* in additional plants will provide valuable information on the origin of monocots. The conservation of this C in monocots would suggest that it plays a role in transcript processing as found for other editing sites in the ndhH-D operon (Del Campo *et al*. 2000, 2002).

The 13th C edited in monocot *ndhC* (Fig. 2*b*) is within a sequence very similar to that surrounding the 7 editing site of barley *ndhB* gene (Freyer *et al*. 1995). Conserved sequences surrounding editing sites, as those of *matK*-2 and *ndhB*-11 editing sites of *Arabidopsis* (Tillich *et al*. 2005), strongly suggest that they act like cis-elements

recognized by the editing machinery.

Knowledge of the editing sites of plant organelle genes will provide valuable evolutionary information and is necessary for proteomic approaches. This requires a complete transcriptome screening of editing sites far beyond the information provided by the genomic sequence of *Marchantia*.

References

Casano, L.M., Zapata, J.M., Martín, M., Sabater, B.: Chlororespiration and poising of cyclic electron transport: plastoquinone as electron transporter between thylakoid NADH dehydrogenase and peroxidase. - J. biol. Chem. **275**: 942-948, 2000.

Casano, L.M., Martín, M., Sabater, B.: Hydrogen peroxide mediates the induction of chloroplast Ndh complex under photooxidative stress in barley. - Plant Physiol. **125**: 1450-1458, 2001.

Cavalier-Smith, T.: Cell and genome co-evolution: facultative anaerobiosis, glycosomes and kinetoplastan RNA editing. - Trends Genet. **13**: 6-9, 1997.

Del Campo, E.M., Sabater, B., Martín, M.: Transcripts of the *ndhH-D* operon of barley plastids: possible role of unedited site III in splicing of the *ndhA* intron. - Nucleic Acids Res. **28**: 1092-1098, 2000.

Del Campo, E.M., Sabater, B., Martín, M.: Post-transcriptional control of chloroplast gene expression: accumulation of stable *psaC* mRNA is due to downstream RNA cleavages in the ndhD gene. - J. biol. Chem. **277**: 36457-36464, 2002.

Del Campo, E.M., Sabater, B., Martín, M.: Characterization of the 5´- and 3´- ends of mRNAs of *ndhH*, *ndhA* and *ndhI* genes of the plastid *ndhH-D* operon. - Biochimie **88**: 347-357, 2006.

Freyer, R., López, C., Maier, R.M., Martín, M., Sabater, B., Kossel, H.: Editing of chloroplast *ndh*B encoded transcript shows divergence between closely related members of the grass family (Poaceae). - Plant mol. Biol. **29**: 679-684, 1995.

Lascano, H.R., Casano, L.M., Martín, M., Sabater, B.: The activity of the chloroplastic Ndh complex is regulated by phosphorylation of the NDH-F subunit. - Plant Physiol. **132**: 256-262, 2003.

Maier, R.M., Zeltz, P., Kössel, H., Bonnard, G., Gualberto, J.M., Grienenberger, J.M.: RNA editing in plant mitochondria and chloroplasts. - Plant mol. Biol. **32**: 343-365, 1996.

Sabater, B., Martín, M., Schmitz-Linneweber, C., Maier, R.M.: Is clustering of plastid RNA editing a consequence of transitory loss of gene function? Implications for past environmental and evolutionary events in plants. - Perspect. Plant Ecol. Evol. System. **5**: 81-89, 2002.

Sazanov, L., Burrows, P.A., Nixon, P.J.: The plastid *ndh* genes code for a NADH-specific dehydrogenase: purification and characterization of a mitochondrial-like complex I from pea thylakoid membranes.- Proc. nat. Acad. Sci. USA **95**: 1319-1324, 1998.

Zapata, J.M., Guéra, A., Esteban-Carrasco, A., Martín, M., Sabater, B.: Chloroplasts regulate leaf senescence: delayed senescence in transgenic *ndhF*-defective tobacco.- Cell Death Differ. **12**: 1277-1284, 2005.