Mol Genet Genomics (2002) 267: 262–269
DOI 10.1007/s00438-002-0662-9

ORIGINAL PAPER

A. Drescher · H. Hupfer · C. Nickel · F. Albertazzi
U. Hohmann · R.G. Herrmann · R.M. Maier

C-to-U conversion in the intercistronic ndhI/*ndhG* RNA of plastids from monocot plants: conventional editing in an unconventional small reading frame?

Received: 15 November 2001 / Accepted: 14 February 2002 / Published online: 26 March 2002
© Springer-Verlag 2002

Abstract Editing of plastid RNAs proceeds by C-to-U, in hornwort species also by extensive U-to-C, transitions, which predominantly lead to the restoration of codons for structurally and/or functionally important, conserved amino acid residues. So far, only one instance of editing outside coding regions has been reported – in the *psbL*/*psbF* intergenic region of *Ginkgo biloba*. This site was proposed to have no functional importance. Here we present an evaluation of an editing site in the *ndhI*/*ndhG* intergenic region in a related group of monocot plants. Efficient editing of this site, as well as the phylogenetic conservation of the resulting uridine residue, point at an important role for the sequence restored by editing. Two potential functions can be envisaged. (1) RNA secondary structure predictions suggest that the C-to-U conversion at this site can lead to a modified stem/loop structure of the *ndhG* 5′ UTR, which could influence *ndhG* expression. (2) Alternatively, editing of the *ndhI*/*ndhG* intergenic region may tag a so far unidentified small (12-codon) ORF, and lead to the restoration of a conserved phenylalanine codon. A screen with specific antibodies elicited against the putative peptide failed to detect such a peptide in chloroplast fractions. However, this failure may be attributable to its low and/or development-specific expression.

Keywords Plastid RNA editing · RNA secondary structure motif · Small peptide-encoding ORF

Communicated by R. Hagemann

A. Drescher · H. Hupfer · C. Nickel · F. Albertazzi[1]
U. Hohmann[2] · R.G. Herrmann · R.M. Maier (✉)
Botanisches Institut der Ludwig-Maximilians-Universität München, Menzingerstr. 67, 80638 München, Germany
E-mail: raimaier@botanik.biologie.uni-muenchen.de
Tel.: +49-89-17861288
Fax: +49-89-1782274

Present addresses:
[1]Centro de Investigacíon en Biología Celular y Molecular, Universidad de Costa Rica, San José, Costa Rica

[2]Institut für Pflanzenbau und Pflanzenzüchtung der Christian-Albrechts-Universität zu Kiel, Olshausenstr. 40, 24098 Kiel, Germany

Introduction

Plastid-encoded transcripts in vascular plants are subject to a substantial degree of post-transcriptional modification based on a variety of processes including RNA editing. With the exception of extensive U-to-C editing in hornwort species (Yoshinaga et al. 1996) plastid RNA editing is generally found as C-to-U changes. Initially, editing in plastids was found to generate translational start codons by ACG to AUG conversions (Hoch et al. 1991; Kudla et al. 1992). Instances of the generation and removal of translational stop codons have also been reported (Wakasugi et al. 1994; Yoshinaga et al. 1996), but usually editing sites reside at internal positions in mRNAs and, with few exceptions, affect the second positions in codons (reviewed in Bock 2000). As a consequence, editing leads to the restoration of codons for conserved amino acid residues, which frequently represent crucial structural and/or functional features of the encoded proteins; this is consistent with the mutant phenotypes or non-functional enzymes that appeared in several cases where editing was impaired (Bock et al. 1994; Zito et al. 1997; Sasaki et al. 2001). The target sequences for the *trans*-acting editing machinery are found to reside in relatively short sequence stretches upstream of editing sites (Bock et al. 1996, 1997a, 1997b; Chaudhuri and Maliga 1996; Reed et al. 2001; Schmitz-Linneweber et al. 2001). However, these *cis*-acting sequences show little general similarity, and nucleus-encoded *trans*-acting factors appear to be diverse as well, with a specific factor for each individual site (Hirose and Sugiura 2001).

Within protein-coding sequences only one case of silent mRNA editing has been reported to date, from UCC(Ser) to UCU(Ser) in tobacco *atpA*, with no apparent change in the coding capacity (Hirose et al. 1996). Similarly only one editing site has been reported to lie

outside of protein-coding regions; this site is found 32 nt upstream of *psbF* in the ancient seed plant *Ginkgo biloba* (Kudla and Bock 1999). This latter site was rated as a non-functional site which has either persisted in the absence of any selective advantage or, alternatively, is incidently recognized by a site-specific factor required for editing at another essential site in the *Ginkgo* plastid chromosome. Plastid RNA editing shares several characteristics with editing in plant mitochondria, which suggests a common evolutionary root for both editing processes. In both organelles, a strong bias for editing of coding regions exists. For instance, in *Arabidopsis thaliana* mitochondria only 7 of the 456 editing sites detected reside in intergenic regions (Giege and Brennicke 1999).

Here we report on C-to-U editing of transcripts of the intercistronic *ndhI*/*ndhG* spacer of plastid DNAs from maize and several other monocot species. Editing at this site produces a highly conserved U residue 10 nt upstream of the translational start codon of *ndhG*, indicating that this site is of functional and/or structural importance.

Materials and methods

Plant material

The plant material used represents eight of the ten putatively monophyletic orders of monocotyledon plant species proposed by the Angiosperm Phylogeny Group (1998). Leaves of *Hydrocharis morsu-ranae* (Alismatales), *Allium cepa*, *Narcissus pseudonarcissus*, *Iris germanica*, *Oncidium ansiferum* (all Asparagales), *Chlorophytum comosum* (Liliales), *Dicranopygium microcephalum* (Pandanales), *Hanguana malayana* (commelinoids), *Tradescantia ohiensis* (Commelinales), *Flagellaria indica*, *Typha angustifolia*, *Juncus inflexus*, *Eriocaulon fenestratum*, *Joinvillea ascendens*, *Eleusine coracana*, *Sorghum bicolor*, *Saccharum officinarum*, *Triticum aestivum*, *Hordeum vulgare* (all Poales), and *Canna indica* (Zingiberales) were kindly provided by the Munich Botanical Garden. Maize (*Zea mays*), oat (*Avena sativa*), rye (*Secale cereale*), *Tripsacum dactyloides*, *Eragrostis tef* and ×*Triticosecale* Wittmack were grown under standard greenhouse conditions. Fresh plant material was homogenized in liquid nitrogen prior to extraction of total nucleic acids.

Isolation of nucleic acids

Total cellular DNA was extracted by the CTAB (cetyltrimethylammonium bromide)-based method (Doyle and Doyle 1990). RNA was isolated using the TRIZOL reagent (GibcoBRL) according to the supplier's instructions. Aliquots of total RNA were treated with DNase I (Roche Diagnostics) for 30 min to remove DNA.

cDNA synthesis and PCR

Prior to reverse transcription DNase I-treated RNA samples were incubated with random hexanucleotide primers at an elevated temperature (65°C) for 5 min and then cooled to room temperature. The reaction was then performed at 37°C for 2 h using SuperScript II reverse transcriptase (GibcoBRL). PCR was done with the following thermal profile: initial denaturation for 3 min at 94°C; 30 cycles of 30 s at 94°C, 30 s at 48°C and 1 min at 72°C, and a final extension step for 3 min at 72°C. The following oligonucleotides were employed for PCR and sequencing in this study: ndhI (5′-TAATTATAGTATTGATTTCGG-3′); ndhG (5′-TAGAATAAGGATATACAAACA-3′); ngseq (labeled with IRD800; 5′-AAGGATATACAAACAAGAAC-3′).

DNA sequencing

Amplified nucleotide sequences were directly determined by the dideoxynucleotide chain-termination method (Sanger et al. 1977) using gene-specific primers. A combination of strategies was employed, including energy-transfer fluorochrome dideoxynucleotide labelling (Rosenblum et al. 1997) with the ABI 377 system (Applied Biosystems), and fluorochrome primer labelling with the LI-COR 4200IR^2 two-laser system (MWG Biotech). For analyses using the ABI system the oligonucleotide ndhG served as the sequencing primer, whereas the fluorescence-labeled primer ngseq was used for the LI-COR system.

Protein analyses

Maize chloroplasts were isolated by isopycnic Percoll gradient centrifugation as described by Robinson and Barnett (1988). Stromal proteins were separated from thylakoids after osmotic shock (for 10 min room temperature in 50 mM HEPES) of chloroplasts and centrifugation for 10 min at 14,000×*g*. Polypeptides were fractionated by SDS-TRIS-glycine-PAGE (Laemmli 1970) in 15% PA containing 7% sucrose, in a gradient gel (12–22% PA and 3–10% sucrose), or by SDS-Tricine-PAGE (Schägger and von Jagow 1987) in 15% PA. For immunoblot analysis, polypeptides were transferred onto either nitrocellulose or polyvinylidene diflouride (PVDF) membranes according to Towbin et al. (1979). The membranes were probed with polyclonal antibodies specific for the "edited" maize plastid-encoded peptide (PEP12) and subsequently incubated with alkaline phosphatase-coupled anti-antibody serum. Alternatively, the enhanced chemoluminescence system (Luminol, Amersham Pharmacia) was used. Antisera were raised against PEP12 in rabbits immunized with synthesized PEP12 peptides coupled either N- or C-terminally to BSA (Eurogenetics).

Multiple alignment

Due to the heterogeneity of intergenic regions, sequences were aligned manually after choosing anchors in the conserved regions of the flanking genes. Publicly available sequences for the entire chloroplast chromosomes of tobacco (*Nicotiana tabacum*), thale cress (*Arabidopsis thaliana*) and rice (*Oryza sativa*), as well as the sequence of the *Synechocystis sp*. PCC6803 genome, were obtained from GenBank. The overall length of the initial alignment was 527 bp. Regions of the multiple alignment with ambiguous homology were omitted from further investigations. The final data matrix contained 324 potentially informative characters, all of which were non-coding.

Phylogenetic analyses

Phylogenetic analyses were performed using the maximum parsimony (MP) method and PAUP version 3.0i. In all analyses the MULPARS, STEEPEST DESCENT and the ACCTRAN options were active. Branches having zero length were collapsed to yield polytomies (COLLAPSE option). The heuristic search option using 100 replicates of simple random sequence addition with Tree Bisection-Reconnection (TBR) branch-swapping was performed to search for multiple islands of most-parsimonious trees (Maddison 1991). All characters have been specified as unordered and unweighted.

Bootstrap analyses (Felsenstein 1985) were used to measure the support of each branch. One hundred bootstrap replications with 10 replications of random sequence addition and TBR branch swapping were performed. In addition, ten thousand randomly generated trees were evaluated to estimate the strength of the phylogenetic signal in a given data set.

Results and discussion

Editing leads to a C-to-U conversion 10 nt upstream of the *ndhG* translation start site in monocot species

The *ndhH* operon of higher plants includes six of the eleven plastid-encoded *ndh* genes (*ndhH*, *A*, *I*, *G*, *E*, *D*) – which specify subunits of an NADH dehydrogenase complex located in the thylakoid membrane (Casano et al. 2000) – together with *psaC*, the gene for subunit VII of Photosystem I. The clustered organization of *ndhA*, *I*, *G* and *E* is highly conserved between plastid chromosomes of plants and even in cyanobacterial genomes (e.g., Kaneko et al. 1996; Kotani and Tabata 1998). The plastid *ndhH* operon is transcribed into a primary polycistronic mRNA, which is then processed to yield various transcript species containing subsets of genes (Matsubayashi et al. 1987; Del Campo et al. 2000; Legen et al. 2001). Comparison between the genomic and cDNA sequences of the *ndhI*/*ndhG* intergenic spacer from maize revealed efficient C-to-U editing at a position 10 nt upstream of the *ndhG* translational start codon (Fig. 1a and b). This site had escaped an initial systematic search for editing sites in the maize plastome because of its unusual position outside of known protein-coding genes (Maier et al. 1995). The occurrence of this intergenic site has been mentioned in a number of reviews on plastid RNA editing (e.g. Maier et al. 1996; Maier 1997; Bock et al. 1997a). Determination of corresponding DNA and cDNA sequences from 27 species representing nearly all of the monocot orders revealed that editing also takes place in the four Poaceae species *Eleusine coracana*, *Saccharum officinarum*, *Tripsacum dactyloides* and *Oryza sativa*, leading to the generation of a U residue that is highly conserved in other monocot species (Fig. 2b and c). A remarkable exception is *Sorghum bicolor*, where the genomic C residue at the editing site remains unchanged. A comparable situation has been reported for one of the *rpoB* sites in maize – which remains unedited in barley (Zeltz et al. 1993). As has been proposed for that site in barley (Zeltz et al. 1993), silencing of the intercistronic editing position in *Sorghum* is possibly tolerated due to compensatory mutations either in *cis* somewhere else in the surrounding sequence (e.g. the A-T exchange at position –35 with respect to the *ndhG* translation start) or in *trans* in gene sequences specifying interacting factors.

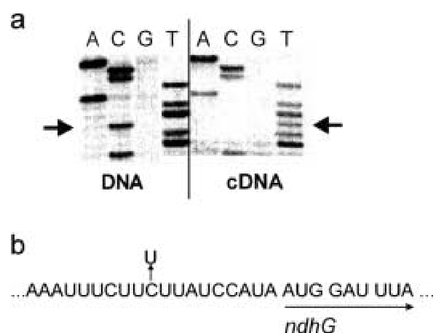

Fig. 1a, b RNA editing in the *ndhI*/*ndhG* intergenic spacer of maize. **a** Autoradiograph showing corresponding DNA and cDNA sequences of the region encompassing the editing site. The position at which editing leads to the C-to-U conversion is marked by an *arrow* in both the DNA and the cDNA sequence. **b** Location of the editing site 10 nt upstream to the translation start codon of maize *ndhG*

The pattern of editing sites frequently differs even between closely related species. This reflects relatively frequent gain and loss of single editing sites during evolution. Available data indicate that the editing site in the *ndhG* 5′ UTR was most probably gained at the branch point leading to Panicoidae (including *Saccharum officinarum*, *Tripsacum dactyloides*, *Zea mays* and *Sorghum bicolor*), Bambusoidae sensu lato (including *Oryza sativa*) and Chloridoidae (including *Eleusine coracana* and *Eragrostis tef*; marked by an asterisk in Fig. 2c) and was subsequently been lost in *Sorghum bicolor* and *Eragrostis tef* (Fig. 2c). An alternative, but less likely, possibility is that the site may have been independently lost in all Pooidae (including ×*Triticosecale* Wittmack, *Triticum aestivum*, *Secale cereale*, *Hordeum vulgare* and *Avena sativa*) and in individual Panicoidae (*Sorghum bicolor*) and Chloridoidae (*Eragrostis tef*) species.

As proposed for the *psbL*/*psbF* intergenic editing site in *Ginkgo* (Kudla and Bock 1999) and mentioned above, editing upstream of *ndhG* may occur accidentally, as a result of adventitious sequence or structural similarities with an essential editing site. To test this assumption, the *ndhG* 5′ UTR sequence was compared with the region around all known plastid editing sites in maize with a window of ±40 bp. No significant similarities were noted. In addition, the complete maize plastome was scanned for sequence similarities to the *ndhG* 5′ UTR in order to find any additional editing site which might have escaped detection in the systematic screen because of its location in a non-coding region. No comparable sequence motif was found elsewhere in the maize plastid chromosome. Thus, if the *ndhI*/*ndhG* intergenic region is incidentally edited by a factor that is required at another essential site, this essential sequence must be located in the mitochondrial or nuclear genome (although no instances of editing have so far been found in nuclear transcripts in plants).

In any case, the high degree of sequence conservation and the crucial role of plastid RNA editing in restoring essential genetic information suggest that the efficiently edited site in the *ndhG* 5′ UTR has an important function. Two mutually non-exclusive alternatives can be considered.

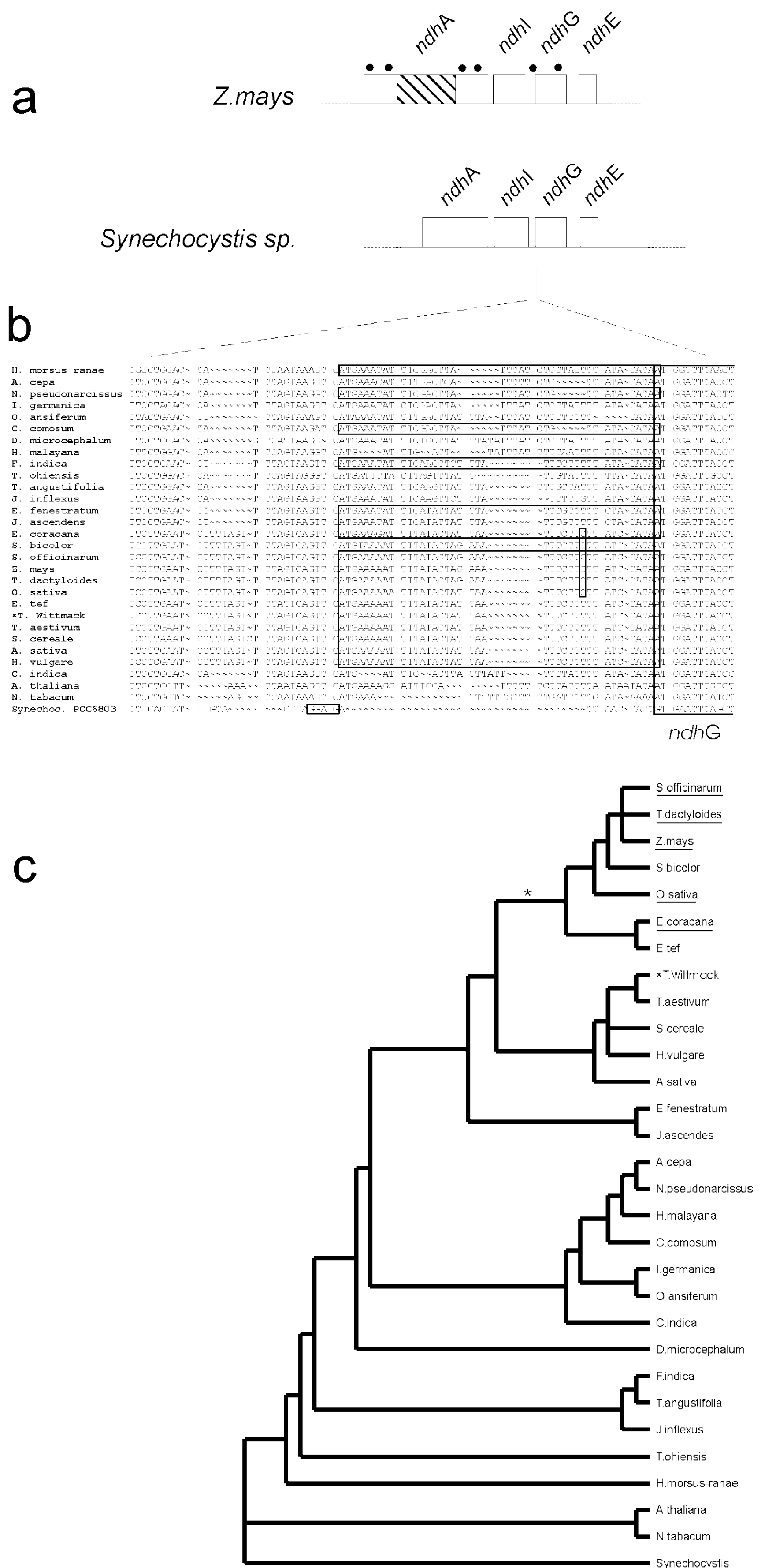

Fig. 2 a Schematic presentation of the conserved *ndhA*, *I*, *G*, *E* gene cluster in maize (*Z. mays*) and cyanobacteria (*Synechocystis* sp.). The group II intron in *ndhA* of maize is marked by the *hatched box*. The six editing positions at which C-to-U transitions post-transcriptionally alter the maize sequences are indicated by *dots*. **b** Multiple sequence alignment of the 5′ UTR of *ndhG*. The C nucleotides that are subject to editing, and the Shine-Dalgarno motif in the *Synechocystis* sequence, are *boxed*. The reading frames homologous to the small ORF in maize are *boxed*. **c** Consensus phylogenetic tree based on Maximum Parsimony (MP) analysis. The search was performed with TBR branch swapping and MULPARS ON, and resulted in twelve equally parsimonious trees. For this search only 100 replications were carried out. Similar results were achieved using 500 replicates with Nearest Neighbor Interchange (NNI) branch swapping. The twelve trees are described by following statistics: Consistency Index (CI) = 0.504, Retention Index (RI) = 0.529, Rescaled Consistency Index (RC) = 0.266. Sequences of *Arabidopsis thaliana*, *Nicotiana tabacum* and of *Synechocystis* PCC 6803 were used as outgroups. The branch at which the editing site in the *ndhG* leader may have been gained is marked by an *asterisk*

Edited and unedited transcripts of the *ndhI/ndhG* intercistronic region can be folded into alternative structures

Since expression control elements, i.e. for translation initiation and transcript processing, are frequently located in 5′ UTRs of plastid genes (e.g. Hirose and Sugiura 1996; Eibl et al. 1999), the C-to-U transition 10 nt upstream of the translation initiation codon could be required to generate a target sequence for (an) interacting factor(s), for instance, by generating an RNA secondary structure essential for *ndhG* expression. To check this, the sequence of the maize *ndhI/ndhG* intergenic region (extending 40 nt to each side of the editing site) was analyzed in its edited and unedited form with the "RNA fold" secondary-structure prediction programme (Mathews et al. 1999). Indeed, the edited RNA sequence can form an energetically less stable alternative secondary structure (Fig. 3). Thus, editing in the *ndhG* 5′ UTR may be necessary to unmask a sequence stretch and make it accessible to interacting *trans* factors (e.g. of the translation machinery). Interestingly, in *Sorghum*, the A/T difference with respect to maize mentioned above could compensate for the lack of editing of the C residue and allow the sequence to fold into the less stable secondary structure which is possible only with the edited RNA in maize (Fig. 3). Thus, if editing is a prerequisite for the maize *ndhG* 5′ UTR to fold into the proposed structure, in *Sorghum* it is not necessary and would even be of disadvantage. In this context, however, the question arises whether the single A/T difference compared to the maize sequence prevents the editing machinery from recognizing/processing the site in *Sorghum* or if a *trans*-acting factor specific for this site was lost with the disappearance of selective pressure for editing. In silico predictions of RNA secondary structures, however, generally depend on the choice of parameters and, therefore, the proposed alternative RNA structure induced by editing needs experimental verification.

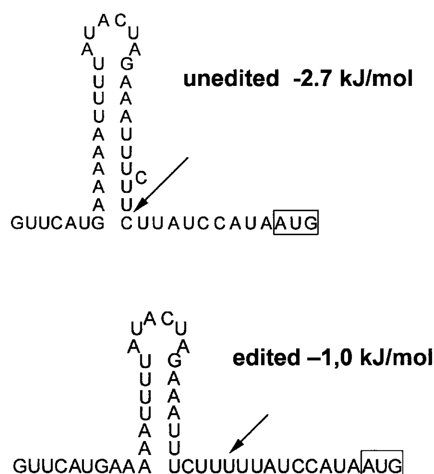

Fig. 3 Predicted alternative RNA secondary structures formed by the unedited and edited *ndhG* 5′ UTRs, calculated using the RNAfold algorithm. The edited and unedited nucleotides are indicated by *arrows*. The *ndhG* translation start codon is *boxed*

RNA editing in the *ndhI/ndhG* intercistronic region may cause a UCU(Ser) to UUU(Phe) conversion in a small open reading frame

An appealing alternative to the proposed role of editing in restoring a functionally required RNA secondary structure in the *ndhI/ndhG* intergenic region is the restoration of a phenylalanine codon by a UCU-to-UUU conversion in an ORF that starts 28 nt upstream of the editing site (Fig. 4a). This putative ORF either may represent a so far unidentified exon or a small gene encoding a peptide comprising only 12 amino acid residues. Short exons, e.g. the 6-bp exon 1 of *petB* or the 8-bp exon 1 of *petD*, are known to exist in plastid chromosomes (Rock et al. 1987; Barkan 1988; Westhoff and Herrmann 1988). Since no sequence motif characteristic of group II exon/intron borders is apparent in the region encoding the ORF, its possible role as an exon that is *trans*-spliced to an exon localized elsewhere in the plastid chromosome is very unlikely.

Small peptides encoded by leader regions are known from bacterial systems. There, so-called "leader peptides" are involved in the regulation of gene expression by transcriptional attenuation (e.g. the 14-residue peptide encoded in the leader of the *trp* operon of *E. coli*; reviewed in Yanofsky 2000). However, the typical RNA

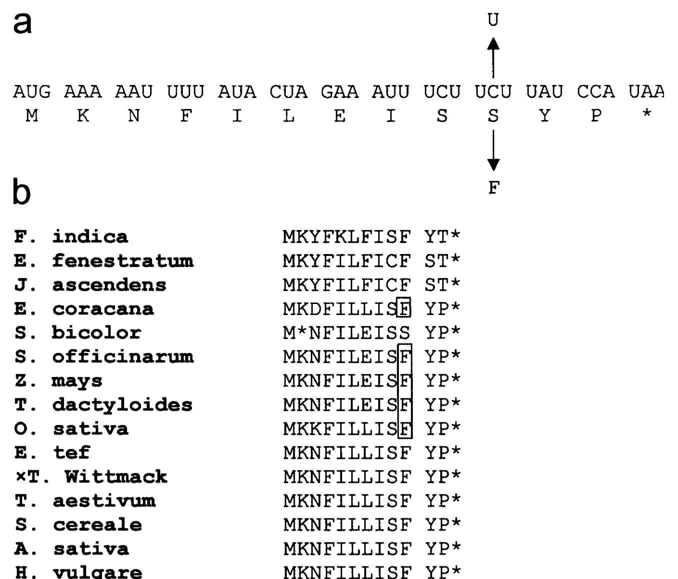

Fig. 4 a Nucleotide and deduced amino acid sequence of the putative peptide-encoding reading frame in the *ndhG* 5′ UTR of maize. C-to-U editing at nucleotide position 29 leads to the conversion of a UCU(Ser) codon to a UUU(Phe) codon. **b** Alignment of amino acid sequences encoded by putative peptide-encoding ORFs in the *ndhG* 5′ UTR of monocot plants. Phenyalanine residues resulting from editing of the respective plastid RNA are *boxed*. Stop codons are marked by *asterisks*. For detailed designation of species see Materials and methods

secondary structure elements involved in the attenuation process in bacteria are not present in the *ndhG* leader region, which makes a role for the encoded peptide in an attenuation process as described for bacteria unlikely.

Several examples of small peptides in plants that are not derived from a larger precursor protein have been described, e.g the decapeptide labaditin in *Jatropha multifida* (Kosasi et al. 1989) or the 12-amino acid residues Enod40 peptide of the legume *Sesbania rostrata* (Corich et al. 1998). Circumstantial evidence suggests that genes coding for short peptides are more widespread than so far supposed. The major problem is to detect these "needles in the haystack" among the enormous background of meaningless short ORFs (Basrai et al. 1997). The editing site in the *ndhI*/*ndhG* intergenic region of maize, however, could operate as a tag for such a small gene. The reading frame is highly conserved in nearly all species of Poaceae examined (Figs. 2 and 4). An interesting exception again is *Sorghum*, which retains its unedited C residue and has no open reading frame here. The ORF is also not found in monocots other than members of the Poaceae, or in dicot species such as tobacco, spinach and *Arabidopsis*. It is known that the coding capacity of plastid chromosomes can differ between species and that genes encoded by plastid chromosomes of one species have been transferred to the nucleus in others. A case in point is *rpl22* (which codes for a polypeptide of the 50S ribosomal subunit) which is found in the plastids in most higher plants but is a nuclear gene in legumes (Gantt et al. 1991). It is therefore conceivable that, in those species where the ORF does not reside in the *ndhI*/*ndhG* intergenic region, it may be located in the nucleus and its products be post-translationally imported into the organelle. However, databank searches using the TBLASTN algorithm (Altschul et al. 1990; with an increased E value of 10,000) did not reveal the presence of a coding gene for a homologous peptide in the current data set including the complete plastid (Sato et al. 1999), mitochondrial (Unseld et al. 1997) and nuclear genomes of *A. thaliana* (Arabidopsis Genome Initiative 2000). Also, no homologs could be found in current EST data resources. This strongly suggests the absence of the peptide in *A. thaliana*. If it exists at all, it is possibly restricted to Poaceae and the closely related Flagellariaceae, Joinvilleaceae and Eriocaulaceae species (Fig. 4b).

To check for expression of the ORF, antibodies were raised against a synthetic peptide (see Materials and methods). The peptide was coupled either N- or C-terminally to BSA, and both versions were used for immunization. The specificity and sensitivity of the antibodies were tested by slot blot analysis. As little as 5 ng of the synthetic peptide was still detectable (Fig. 5a). However, no signal could be found with total cellular proteins of maize leaves and soluble or membrane fractions of maize chloroplasts, even when preparative high-resolution SDS-PA gels loaded with up to 50 μg of protein sample (Fig. 5b) were probed. Thus, if synthesized, the putative peptide must accumulate only

Fig. 5 a Slot-blot immunoanalysis: Different amounts of synthetically synthesized peptide (PEP12) and BSA, respectively, were transferred to a PVDF membrane and probed with an antiserum raised against the BSA-coupled synthetic peptide. **b** Western analysis. Aliquots of PEP12 (5 μg) and protein samples (50 μg) isolated from soluble and membrane fractions, respectively, of maize chloroplasts were fractionated electrophoretically in a 12–22% gradient polyacrylamide gel. Proteins were then blotted onto a PVDF membrane and probed with the PEP12-specific antiserum

in very low amounts, below the detection limit of approximately 1 ng of protein attainable with the Enhanced Chemiluminescence system (Luminol, Amersham Pharmacia). This would be expected, if a role in regulation (for instance in signal transduction) is assumed. For the isolation of systemin, the first plant peptide messenger detected (McGurl et al. 1992), 28 kg of tomato leaves was required, and 30,000 tomato plants had to be processed to produce 1 μg of the peptide (Marx 1996). On the other hand, expression of the proposed peptide may be restricted to certain tissues and/or developmental stages. Further work will therefore be needed to settle the question of whether the putative peptide exists or not.

Acknowledgements This work was supported by DFG (MA 1592 and SFB184). F.A. was a recipient of a fellowship from the Deutscher Akademischer Austauschdienst (DAAD). We wish to thank Christian Schmitz-Linneweber for helpful discussions.

References

Altschul FA, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Angiosperm Phylogeny Group (APG) (1998) An ordinal classification for the families of flowering plants. Ann Miss Bot Garden 80:531–553

Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408:796–815

Barkan A (1988) Proteins encoded by a complex chloroplast transcription unit are each translated from both monocistronic and polycistronic mRNAs. EMBO J 7:2637–2644

Basrai MA, Hieter P, Boecke JD (1997) Small open reading frames: beautiful needles in the haystack. Genome Res 7:768–771

Bock R (2000) Sense from nonsense: how the genetic information of chloroplasts is altered by RNA editing. Biochimie 82:549–557

Bock R, Kössel H, Maliga P (1994) Introduction of a heterologous editing site into the tobacco plastid genome: the lack of RNA editing leads to a mutant phenotype. EMBO J 13:4623–4628

Bock R, Hermann M, Kössel H (1996) In vivo dissection of *cis*-acting determinants for plastid RNA editing. EMBO J 15: 5052–5059

Bock R, Albertazzi F, Freyer R, Fuchs M, Ruf S, Zeltz P, Maier RM (1997a) Transcript editing in chloroplasts of higher plants. In: Schenk HEA, Herrmann RG, Jeon KW, Müller NE, Schwemmler W (eds) Eukaryotism and symbiosis. Springer, Berlin-Heidelberg-New York, pp 123–137

Bock R, Hermann M, Fuchs M (1997b) Identification of critical nucleotide positions for plastid RNA editing site recognition. RNA 3:1194–1200

Casano LM, Zapata JM, Martin M, Sabater B (2000) Chlororespiration and poising of cyclic electron transport. J Biol Chem 275:942–948

Chaudhuri S, Maliga P (1996) Sequences directing C to U editing of the plastid *psbL* mRNA are located within a 22 nucleotide segment spanning the editing site. EMBO J 15:5958–5964

Corich V, Goormachtig S, Lievens S, Van Montagu M, Holsters M (1998) Patterns of ENOD40 gene expression in stem-borne nodules of *Sesbania rostrata*. Plant Mol Biol 37:67–76

Del Campo EM, Sabater B, Martin M (2000) 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

Doyle JJ, Doyle L (1990) Isolation of plant DNA from fresh tissue. Focus 12:13–15

Eibl C, Zou Z, Beck A, Kim M, Mullet J, Koop HU (1999) In vivo analysis of plastid *psbA*, *rbcL* and *rpl32* UTR elements by chloroplast transformation: tobacco plastid gene expression is controlled by modulation of transcript levels and translation efficiency. Plant J 19:333–345

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791

Gantt JS, Baldauf SL, Calie PJ, Weeden NF, Palmer JD (1991) Transfer of *rpl22* to the nucleus greatly preceded its loss from the chloroplast and involved the gain of an intron. EMBO J 10:3073–3078

Giege P, Brennicke A (1999) RNA editing in *Arabidopsis* mitochondria effects 441 C to U changes in ORFs. Proc Natl Acad Sci USA 96:15324–15329

Hirose T, Sugiura M (1996) *Cis*-acting elements and *trans*-acting factors for accurate translation of chloroplast *psbA* mRNAs: development of an in vitro translation system from tobacco chloroplasts. EMBO J 15:1687–1695

Hirose T, Sugiura M (2001) Involvement of a site-specific *trans*-acting factor and a common RNA-binding protein in the editing of chloroplast mRNAs: development of a chloroplast in vitro RNA editing system. EMBO J 20:1144–1152

Hirose T, Fan H, Suzuki JY, Wakasugi T, Tsudzuki T, Kössel H, Sugiura M (1996) Occurrence of silent RNA editing in chloroplasts: its species specificity and the influence of environmental and developmental conditions. Plant Mol Biol 30:667–672

Hoch B, Maier RM, Appel K, Igloi G L, Kössel H (1991) Editing of a chloroplast mRNA by creation of an initiation codon. Nature 353:178–180

Kaneko T, et al (1996) Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis sp.* strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. DNA Res 3:109–136

Kosasi S, Van der Sluis WG, Boelens RT, Hart LA, Labadie RP (1989) Labaditin, a novel cyclic decapeptide from the latex of *Jatropha multifida* L. (Euphorbiaceae). Isolation and sequence determination by means of two-dimensional NMR. FEBS Lett 256:91–96

Kotani H, Tabata S (1998) Lessons from sequencing of the genome of a unicellular cyanobacterium, *Synechocystis* sp. PCC6803. Annu Rev Plant Physiol 49:151–171

Kudla J, Bock R (1999) RNA editing in an untranslated region of the *Ginkgo* chloroplast genome. Gene 234:81–86

Kudla J, Igloi GL, Metzlaff M, Hagemann R, Kössel H (1992) RNA editing in tobacco chloroplasts leads to the formation of a translatable *psbL* mRNA by a C to U substitution within the initiation codon. EMBO J 11:1099–1103

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680–685

Legen J, Schmitz-Linneweber C, Drescher A, Hupfer H, Tillich M, Herrmann RG, Maier RM (2001) Decoding of the *ndhH*-operon from spinach: an example for the complexity of plastid gene expression in higher plants. Endocytobiosis Cell Res 14:11–19

Maddison DR (1991) Discovery and importance of multiple islands of most-parsimonious trees. Syst Zool 40:315–328

Maier RM (1997) Editing of plastid encoded transcripts. Endocytobiosis Cell Res 12:41–48

Maier RM, Neckermann K, Igloi GL, Kössel H (1995) Complete sequence of the maize chloroplast genome: gene content, hotspots of divergence and fine tuning of genetic information by transcript editing. J Mol Biol 251:614–621

Maier RM, Zeltz P, Kössel H, Bonnard G, Gualberto JM, Grienenberger JM (1996) RNA editing in plant mitochondria and chloroplasts. Plant Mol Biol 32:343–365

Marx J (1996) Plants, like animals, may make use of peptide signals. Science 273:1338–1339

Mathews DH, Sabina J, Zuker M, Turner DH (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. J Mol Biol 288:911–940

Matsubayashi T, Wakasugi T, Shinozaki K, Yamagushi-Shinozaki K, Zaita N, Hidaka T, Meng BY, Ohto C, Tanaka M, Kato A, Maruyama T, Sugiura M (1987) Six chloroplast genes (*ndhA-F*) homologous to human mitochondrial genes encoding components of the respiratory chain NADH dehydrogenase are actively expressed: determination of splice sites in *ndhA* and *ndhB* pre-mRNAs. Mol Gen Genet 210:385–393

McGurl B, Pearce G, Orozco-Cardenas M, Ryan CA (1992) Structure, expression, and antisense inhibition of the systemin precursor gene. Science 255:1570–1573

Reed ML, Peeters NM, Hanson MR (2001) A single alteration 20 nt 5′ to an editing target inhibits chloroplast RNA editing in vivo. Nucleic Acids Res 29:1507–1513

Robinson C, Barnett LK (1988) Isolation and analysis of chloroplasts. In: Shaw CH (ed) Plant molecular biology: a practical approach. IRL Press Oxford, pp 67–78

Rock CD, Barkan A, Taylor WC (1987) The maize plastid *psbB-psbF-petB-petD* gene cluster: spliced and unspliced *petB* and *petD* RNAs encode alternative products. Curr Genet 12:69–77

Rosenblum BB, Lee LG, Spurgeon SL, Khan SH, Menchen SM, Heiner CR, Chen SM (1997) New dye-labeled terminators for improved DNA sequencing patterns. Nucleic Acids Res 25:4500–4504

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467

Sasaki Y, Kozaki A, Ohmori A, Iguchi H, Nagano Y (2001) Chloroplast RNA editing required for functional acetyl-CoA carboxylase in plants. J Biol Chem 276:3937–3940

Sato S, Nakamura Y, Kaneko T, Asamizu E, Tabata S (1999) Complete structure of the chloroplast genome of *Arabidopsis thaliana*. DNA Res 6:283–290

Schägger H, von Jagow G (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. Anal Biochem 166:368–379

Schmitz-Linneweber C, Tillich M, Herrmann RG, Maier RM (2001) Heterologous, splicing dependent RNA editing in

chloroplasts: allotetraploidy provides transfactors. EMBO J 20:4874–4883

Towbin H, Staehelin T, Gordon J (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc Natl Acad Sci USA 76:4350–4354

Unseld M, Marienfeld JR, Brandt P, Brennicke A (1997) The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides. Nat Genet 15:57–61

Wakasugi T, Hirose T, Horihata M, Tsudzuki T, Kossel H, Sugiura M (1996) Creation of a novel protein-coding region at the RNA level in black pine chloroplasts: the pattern of RNA editing in the gymnosperm chloroplast is different from that in angiosperms. Proc Natl Acad Sci USA 93:8766–8770

Westhoff P, Herrmann RG (1988) Complex RNA maturation in chloroplasts. The *psbB* operon from spinach. Eur J Biochem 171:551–564

Yanofsky C (2000) Transcript anttenuation: once viewed as a novel regulatory strategy. J Bacteriol 182:1–8

Yoshinaga K, Iinuma H, Masuzawa T, Uedal K (1996) Extensive RNA editing of U to C in addition to C to U substitution in the *rbcL* transcripts of hornwort chloroplasts and the origin of RNA editing in green plants. Nucleic Acids Res 24:1008–1014

Zeltz P, Hess WR, Neckermann K, Börner T, Kössel H (1993) Editing of the chloroplast *rpoB* transcript is independent of chloroplast translation and shows different patterns in barley and maize. EMBO J 12:4291–4296

Zito F, Kuras R, Choquet Y, Kössel H, Wollman FA (1997) Mutations of cytochrome b_6 in *Chlamydomonas reinhardtii* disclose the functional significance for a proline to leucine conversion by *petB* editing in maize and tobacco. Plant Mol Biol 33:79–86