Substitutional editing of transcripts from genes of cyanobacterial origin in the dinoflagellate *Ceratium horridum*☆

Stefan Zauner[a,*], Doris Greilinger[a], Thomas Laatsch[a], Klaus V. Kowallik[b], Uwe-G. Maier[a]

[a] *Cell Biology, Philipps-University Marburg, Karl-von-Frisch Strasse, D-35032 Marburg, Germany*
[b] *Institute of Botany I, Heinrich-Heine-University Düsseldorf, Universitätsstrasse 1, D-40225 Düsseldorf, Germany*

Received 30 August 2004; revised 20 October 2004; accepted 21 October 2004

Available online 3 November 2004

Edited by Horst Feldmann

Abstract Peridinin-containing dinoflagellates, a group of alveolate organisms, harbour small plasmids called minicircles. As most of these minicircles encode genes of cyanobacterial origin, which are also found in plastid genomes of stramenopiles, they were thought to represent the plastid genome of peridinin-containing dinoflagellates. The analyses of minicircle derived mRNAs and the 16S rRNA showed that extensive editing of minicircle gene transcripts is common for *Ceratium horridum*. Posttranscriptional changes occur predominantly by editing A into G, but other types of editing including a previously unreported A to C transversion were also detected. This leads to amino acid changes in most cases or, in one case, to the elimination of a stop-codon. Interestingly, the edited mRNAs show higher identities to homologous sequences of other peridinin-containing dinoflagellates than their genomic copy. Thus, our results imply that transcript editing of genes of cyanobacterial origin is species specific in peridinin-containing dinoflagellates and demonstrate that editing of genes of cyanobacterial origin is not restricted to land plants.

Keywords: Dinoflagellate; Minicircle; RNA editing; Transition; *Ceratium horridum*

1. Introduction

Dinoflagellates are unusual unicellular eukaryotes [1–7]. The spontaneous occurrence of red tides and the sensitive symbiosis with corals highlighted the dinoflagellates important role in ecosystem stability [8]. In the heterogeneous group of dinoflagellates about 50% are phototrophs, most of them contain chlorophyll c_2 and peridinin as the main carotenoids. Dinoflagellates possess characteristics that are unique for eukaryotic organisms. Examples for this are the high amount of permanently condensed nuclear DNA in most species [9,10], the lack of histone proteins and consequently of nucleosomes [11–13], proteins encoded as polyproteins [14,15], and, in the case of peridinin-containing dinoflagellates, a proteobacterial-like nuclear encoded type II RuBisCO [16,17]. As recently shown, dinoflagellates are furthermore the ''champions'' in successfully transferring genes from the plastid into the cell nucleus [18,19]. Whether the plastids of peridinin-containing dinoflagellates, surrounded by an envelope of three membranes, evolved through secondary or tertiary endosymbiosis remains controversial to date [7,20,21].

Plastid genomes of most eukaryotes are circular molecules 100–200 kb in length [22,23]. This is thought not to be the case in peridinin-containing dinoflagellates such as *Heterocapsa*, *Amphidinium* and *Symbiodinium*, which harbour minicircles encoding one or sometimes two genes. ''Empty minicircles'' without gene sequences are also encountered [1–6,24].

Several database entries of minicircle-encoded genes show that the coding regions are not preceded by an initiator methionine codon. As dinoflagellates are furthermore known to edit mitochondrial genes [25], *A. operculatum* was studied to assess whether an initiator codon was posttranscriptionally created by RNA editing of minicircle-derived transcripts. These experiments revealed that this is not the case [5] and RNA editing of genes of cyanobacterial origin was thus considered to be restricted to higher plants [26,27].

2. Materials and methods

2.1. Data deposition

GenBank Accession Nos.: AJ628833–AJ628838, AF490362, and AF490358.

2.2. Strain

C. horridum was obtained and cultivated as described [28].

2.3. RNA isolation and RT-PCR

RNA was isolated using the peqGold RNAPure FL isolation kit (peqLab Biotechnologie GmbH) following the manufacturer's instructions. RT-PCRs were performed using the cMaster™ RT_{plus}PCR System (Eppendorf). The purity of the RNA preparation was checked by PCR without reverse transcriptase. The primers are listed in Supplement Table 1.

2.4. Cloning and sequencing

RT-PCR products were ligated into pGEM®-T vektor (Promega), amplified in *Escherichia coli* MRF' and sequenced on a Li-Cor 4200 sequencer (MWG-Biotech) using IRD700/IRD800 labelled M13-20 standard primers and the 'Thermosequenase Fluorescent Labelled Primer Cycle Sequencing Kit with 7-deaza-dGTP' (Amersham). At least three clones of independent PCRs were sequenced. Alignments were assembled using Sequencher (GeneCodes).

3. Results and discussion

We have isolated minicircles from the dinoflagellate *Ceratium horridum* [28], which resemble those of other

☆ Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.febslet.2004.10.060.

* Corresponding author. Fax: +49-6421-2822057.
E-mail address: zauner@staff.uni-marburg.de (S. Zauner).

0014-5793/$22.00 © 2004 Published by Elsevier B.V. on behalf of the Federation of European Biochemical Societies.
doi:10.1016/j.febslet.2004.10.060

dinoflagellates [1–6]: minicircles are small replicons, usually encoding one, rarely two genes of cyanobacterial origin, and sometimes no genes. When possible, we isolated minicircle genes from total DNA for cloning instead of amplifying them by PCR. This Protocol succeeded in isolating a minicircle encoding a *psb*B-gene, in which the genomic sequence of the potential coding region is interrupted by a frameshift caused by a 17 bp insertion. This finding prompted us to test if a second *psb*B-encoding minicircle with a perfect coding region exists in *C. horridum*. We thus amplified *psb*B genes using different primer pairs surrounding the 17 bp insertion and isolated a second *psb*B-gene which is identical to the genomic clone, but lacks the 17 bp insertion. Thus, two *psb*B genes exist in *C. horridum*, which are identical in their sequence except for the 17 bp insertion.

RT-PCR experiments were used to test the *psb*B transcripts. These showed two different transcripts in high resolution polyacrylamide gels, one identical in length to the insertion harbouring genomic copy and the second with a smaller size. Sequencing of the cloned RT-products indicated that the smaller amplified product represents either the transcript of the insertion-less *psb*B-gene or a spliced product of the insertion-harbouring gene, whereas the longer transcript was identical to the genomic copy containing the 17 bp insertion. To our surprise, the mature insertion-less *psb*B transcript differs from its respective DNA sequence by extensive editing (Table 1). We were not able to identify this edited sequence on the DNA level using PCR with primers derived from unedited parts of the sequences. From 1553 bases analysed, the *psb*B transcript is edited at 86 positions, corresponding to 5.5% of nucleotide substitutions. Predominantly, A to G transitions were detected (32 positions), followed by 29 editing events changing U to C (Table 1). Remarkably, two further transversions were identified: G to C, which is only known for mitochondrial mRNAs in some dinoflagellates [25] and A to C, which was previously unreported for any edited RNA. Editing was found at all three codon positions leading to 76 amino acid and 6 silent changes (Table 1). Interestingly, 34 editing events increase the amino acid sequence identity of the edited transcript of *psb*B with its homologous gene of *Heterocapsa triquetra* (Fig. 1(a)). In contrast, only 7 amino acid changes caused by editing impair the percentage rate of identity.

To characterize editing of cyanobacterial-derived gene transcripts in *C. horridum* in more detail, we included three further minicircle-encoded genes and their transcripts in our analyses: *psb*E, *psa*A, and a part of a putative 16S rRNA gene of *C. horridum*. In the case of the *psb*E gene and transcript, we identified a genomic sequence and its corresponding mRNA which are identical in length. From this small gene, 171 bases were analysed for RNA editing. We found changes in the transcript at 13 positions (7.6%) when compared with the genomic counterpart. Substitutional editing led to 10 amino acid changes, one of the editing events being silent. The distribution of the type of editing is more balanced than in the *psb*B-minicircle (Table 1), though A to G as well as G to C changes were also observed.

The analysis of minicircle-encoded *psa*A genes of *C. horridum* revealed two sequences that differ by a 24 bp deletion. RT-PCR showed that only the gene containing the 24 bp insertion version is transcribed. Thus, we consider the gene with the deletion a non-transcribed pseudogene. In a total of 1279 bases of the *psa*A gene analysed, we detected 104 changes caused by substitutional editing of the mRNA (8.1%). Of these, approximately 40% (41 events) are transitions from A to G followed by 28 editing events (27%) changing U to C, a distribution similar to that of the *psb*B transcript. As in the mature *psb*B transcript, we also found G to C and A to C transversions in the mature *psa*A transcript, though to a lower extent (Table 1).

As shown in Fig. 1(b), editing of *psa*A transforms a stop-codon into cystein and increases the identity to a homologous sequence of *H. triquetra*. 34 editing events increase the identity of encoded amino acids between the mRNA of *C. horridum* and the genomic sequence of *H. triquetra*, 8 changes by editing decrease the identity.

Editing also occurs in a minicircle with homology to parts of a sequence annotated as the rrs gene of *H. triquetra*. By analysing the transcript of this minicircle region, we identified 35 substitutional changes (3.6%) in 976 sites investigated (Table 1). The changes occur predominantly in transitions from A to G (77%). However, as shown in Table 1, this RNA has one remarkable editing event that changes U to G, which was not detected in the mRNAs of minicircle genes investigated thus far. On the other hand, editing changing G to C, which prior to our findings was only described for transcripts of mitochondrial origin of some dinoflagellates [25], was not detected in the putative *rrs* gene transcript. Again, contrary to the *psb*B and *psa*A minicircle transcripts, editing of the putative *rrs* gene decreases the similarity to the *H. triquetra* sequence significantly (decrease in 14, increase in 9 positions). Regardless of our doubts as to whether the *H. triquetra* sequence and therefore also the *C. horridum* homologue indeed represent functional *rrs* genes, the noted differences to mRNAs in distribution and quality of editing could reflect the structural needs of rRNAs, in which the secondary/tertiary structures and the interactions with ribosomal proteins, but not

Table 1
Types of substitutions inferred in the *C. horridum* minicircle genes *psa*A, *psb*B, *psb*E and a putative *rrs* gene

	*psa*A	*psb*B	*psb*E	1st	2nd	3rd	16S rRNA
A → G	41	32	2	57	15	3	27
G → A	11	10	3	18	2	4	–
C → U	12	9	2	17	2	4	6
U → C	28	29	5	27	30	5	1
G → C	9	5	1	3	10	2	–
A → C	3	1	–	2	2	–	–
U → G	–	–	–	–	–	–	1
% of sites edited	*8.1*	*5.5*	*7.6*	–	–	–	*3.6*

Indicated is the type of substitution, its frequencies in the mRNAs and at the bottom the percentage of sites edited in the genes. 1st, 2nd, and 3rd indicate the edited codon position within the codons of *psa*A, *psb*B and *psb*E.

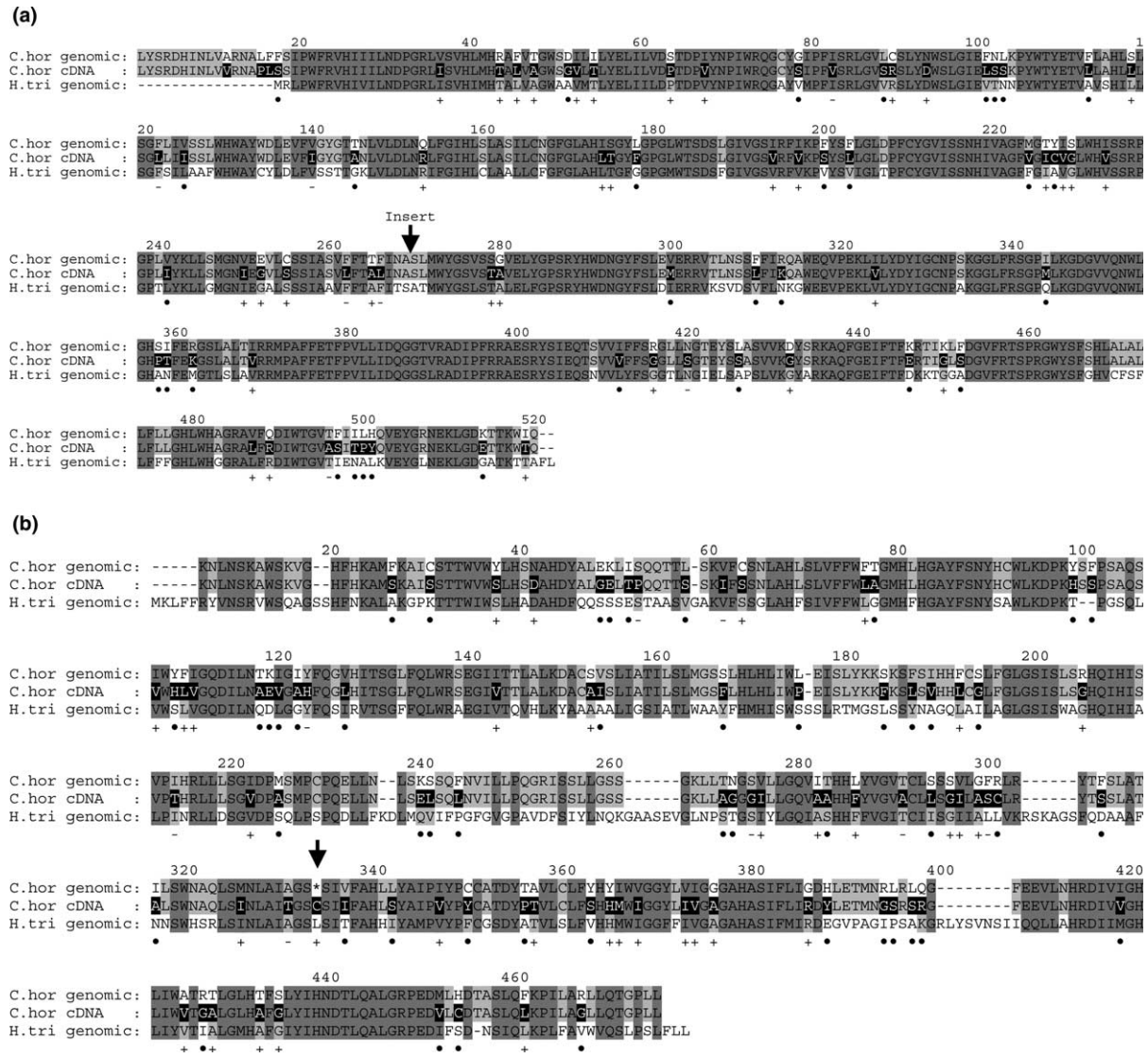

Fig. 1. Alignments of the amino acid-sequences derived from genomic sequences and the mRNAs of (a) *psb*B and (b) *psa*A of *C. horridum* with the homologous sequences from *H. triquetra* (GenBank Accession No. *psa*A: AAD44698; *psb*B: AAD44701). "–" indicates decrease in identity, "+" indicates increase in identity. "·" indicates editing events that neither increase nor decrease the degree of identity. (a) Arrow indicates the position of the 17bp insertion. (b) The in-frame stop codon which is eliminated by substitutional editing is indicated by an arrowhead.

codons or codon usage have to be considered by posttranscriptional changes. As this putative rRNA gene is not completely isolated, we were not able to predict any impact of editing in respect to the secondary structure of this RNA.

Peridinin-containing dinoflagellates are not only enigmatic for their genomic organization and the use of a type II RuBisCO, but also because of their unusual minicircles and the high amount of nucleus-encoded genes for plastid functions [18,19]. Our results expand this list of aberrant characters of peridinin-containing dinoflagellates by the detection of RNA editing of cyanobacterial-derived genes, which prior to our findings was exclusively reserved for land plants [26]. Interestingly, the editing mechanisms lead to a significantly higher number of presumed A to I nucleotide changes, which were identified as an A to G difference between genomic sequence and edited transcript [29]. In eukaryotes this type of editing is known as nucleus-specific, causing substitutions in tRNAs from yeast and mRNAs from nucleus-encoded genes [29]. The only known A to G editing of organellar transcripts is in mitochondrial mRNAs of some dinoflagellates [25], where A to G changes as well as the other here reported editing events occur in a similar ratio to the minicircle genes of *C. horridum*. The exceptions are the previously unreported A to C transversions and the frequent occurrence of G to A conversion in minicircle encoded transcripts. This leads to the prediction of the coexistence of a wide variety of editing mechanisms, some of them still not known. However, the editing sites in mitochondrial transcripts of dinoflagellates are clustered outside the conserved regions [25]. It was therefore speculated that the nucleus-specific A to G conversion in the mitochondrial transcripts could be introduced by a guide RNA [25] and not by a recently characterized family of adenosine deaminases, which are responsible for catalysing A to G conversions in nuclear pre-mRNAs [30–32]. We did not observe a clustering

of editing sites in the transcripts of minicircle-encoded genes, suggesting that adenosine deaminases are acting on minicircle transcripts. This may confirm our previous findings that the minicircles of *C. horridum* are nucleus localized [28]. However, the question of whether the transcripts of minicircle-genes mature in the cell nucleus or whether the RNAs are imported into the plastid, in which an editing machinery similar to that of the mitochondrion [25] is working, has yet to be answered.

As mentioned above, editing of mRNAs of *C. horridum* minicircles increases the identity of the encoded protein to the homologous sequences of *H. triquetra* or *A. operculatum*. This finding, together with the fact that the genes for *psb*B as well as *psa*A and *atp*B from *A. operculatum* show no evidence for RNA editing [5], implies that RNA editing is species specific in dinoflagellates.

Peridinin-containing dinoflagellates do not seem to be a consistent group with respect to their minicircles. At least *C. horridum* differs significantly from other dinoflagellates by editing of genes of cyanobacterial origin.

Acknowledgements: We thank Sven Gould, Bettina Stoebe-Maier and Rainer Maier for comments and suggestions. Our work is supported by the Deutsche Forschungsgemeinschaft (SFB-TR1). We dedicate this paper to Prof. Dr. Rainer Maier, who passed away too early in April 2004.

References

[1] Zhang, Z., Green, B.R. and Cavalier-Smith, T. (1999) Nature 400, 155–159.
[2] Zhang, Z., Cavalier-Smith, T. and Green, B.R. (2001) Mol. Biol. Evol. 18, 1558–1565.
[3] Zhang, Z., Cavalier-Smith, T. and Green, B.R. (2002) Mol. Biol. Evol. 19, 489–500.
[4] Barbrook, A.C. and Howe, C.J. (2000) Mol. Gen. Genet. 263, 152–158.
[5] Barbrook, A.C., Symington, H., Nisbet, R.E., Larkum, A. and Howe, C.J. (2001) Mol. Genet. Genomics 266, 632–638.
[6] Hiller, R.G. (2001) FEBS Lett. 505, 449–452.
[7] Stoebe, B. and Maier, U.-G. (2002) Protoplasma 219, 123–130.
[8] Schnepf, E. and Elbraechter, M. (1999) Grana 38, 81–97.
[9] Soyer-Gobillard, M.O., Gillet, B., Geraud, M.L. and Bhaud, Y. (1999) Int. Microbiol. 2, 93–102.
[10] Bhaud, Y., Guillebault, D., Lennon, J.F., Defacque, H., Soyer-Golbillard, M.O. and Moreau, H. (2000) J. Cell. Sci. 113, 1231–1239.
[11] Herzog, M. and Soyer, M.O. (1983) Eur. J. Cell Biol. 30, 33–41.
[12] Rizzo, P.J. (1987) in: The Biology of Dinoflagellates (Taylor, F.J.R., Ed.), pp. 143–173, Blackwell Botanical Monographs.
[13] Kasinsky, H.E., Lewis, J.D., Dacks, J.B. and Ausio, J. (2001) FASEB J. 15, 34–42.
[14] Hiller, R.G., Wrench, P.M. and Sharples, F.P. (1995) FEBS Lett. 363, 175–178.
[15] Rowan, R., Whitney, S.M., Fowler, A. and Yellowlees, D. (1996) Plant Cell 8, 543–545.
[16] Morse, D., Salois, P., Markovic, P. and Hastings, J.W. (1995) Science 268, 1622–1624.
[17] Whitney, S.M., Shaw, D.C. and Yellowlees, D. (1995) Proc. R. Soc. Lond. B Biol. Sci. 259, 271–275.
[18] Hackett, J.D., Yoon, H.S., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Nosenko, T. and Bhattacharya, D. (2004) Curr. Biol. 14, 213–218.
[19] Bachvaroff, T.R., Concepcion, G.T., Rogers, C.R., Herman, E.M. and Delwiche, C.F. (2004) Protist 155, 65–78.
[20] Saldarriaga, J.F., Taylor, F.J., Keeling, P.J. and Cavalier-Smith, T. (2001) J. Mol. Evol. 53, 204–213.
[21] Yoon, H.S., Hackett, J.D. and Bhattacharya, D. (2002) Proc. Natl. Acad. Sci. USA 99, 11724–11729.
[22] Martin, W., Rujan, T., Richly, E., Hansen, A., Cornelsen, S., Lins, T., Leister, D., Stoebe, B., Hasegawa, M. and Penny, D. (2002) Proc. Natl. Acad. Sci. USA 99, 12246–12251.
[23] Koumandou, V.L., Nisbet, R.E.R., Barbrook, A.C. and Howe, C.J. (2004) Trends Genet. 20, 261–267.
[24] Takishita, K., Ishikura, M., Koike, K. and Maruyama, T. (2003) Phycologia 42, 285–291.
[25] Lin, S., Zhang, H., Spencer, D.F., Norman, J.E. and Gray, M.W. (2002) J. Mol. Biol. 320, 727–739.
[26] Sabater, B., Martín, M., Schmitz-Linneweber, C. and Maier, R.M. (2002) Perspect Plant Ecol., Evol Systemat. 5, 81–90.
[27] Herrmann, R.G., Maier, R.M. and Schmitz-Linneweber, C. (2003) Philos. Trans. R. Soc. Lond. B Biol. Sci. 358, 87–97.
[28] Laatsch, T., Zauner, S., Stoebe-Maier, B., Kowallik, K.V. and Maier, U.-G. (2004) Mol. Biol. Evol. 21, 1318–1322.
[29] Gott, J.M. and Emeson, R.B. (2000) Annu. Rev. Genet. 34, 499–531.
[30] Maas, S., Melcher, T. and Seeburg, P.H. (1997) Curr. Opin. Cell Biol. 9, 343–349.
[31] Gerber, A. and Keller, W. (2001) Trends Biochem. Sci. 26, 376–384.
[32] Bass, B.L. (2002) Annu. Rev. Biochem. 71, 817–846.