Volume 221, number 2, 189–193 FEB 05076 September 1987

*Commentary–Hypothesis*

# Unusual 5 S ribosomal RNAs

## An analysis of individual segments can reveal phylogenetic relatedness

Nicolas Delihas

*Department of Microbiology, School of Medicine, SUNY, Stony Brook, NY 11794, USA*

Received 10 June 1987

Sequence comparisons of 5 S and other ribosomal RNAs by segments can be useful in understanding anomalous primary and secondary structures and in assessing phylogenetic relationships. In a segmented analysis, the 5'-half of the *Chlamydomonas reinhardii* chloroplast 5 S ribosomal RNA is found to have a very close sequence homology to the green plant chloroplast and cyanobacterial 5 S RNAs; however, the 3'-half has a highly unusual sequence. Further comparisons of homologies between regions of the 5 S RNAs from *C. reinhardii* and the green plant chloroplasts suggest that genetic rearrangements within the 5 S DNA may have produced the unusual sequence at the 3'-half. Segmented analyses of the *C. reinhardii* and green plant chloroplast 5 S RNAs suggest a close relationship which is not revealed by overall sequence comparisons.

5 S ribosomal RNA; Segmented analysis; Gene rearrangement; Phylogenetics; Chloroplast origin

## 1. INTRODUCTION

Over 500 5 S ribosomal RNA sequences are known and a comparison of these sequences reveals that the 5 S RNA genes have been highly conserved during evolution and that all sequences conform to a generalized 5 S RNA secondary structural model. Thus, an analysis of 5 S RNA primary structure as well as group-specific signatures (i.e. nucleotide insertions, deletions or substitutions and secondary structural features that are characteristic of a particular phylogenetic group) have been used to ascertain phylogenetic relationships between organisms [1]. In some cases, when suspected genetic rearrangements within the ribosomal structural genes have occurred, an overall sequence analysis by residue-to-residue comparison may not yield meaningful phylogenetic relationships. A comparison of 5 S RNA sequence homology by individual segments can provide information on suspected genetic rearrangements. In this paper we show how possible gene rearrangements can be accounted for in analyzing 5 S RNA secondary structure and in 5 S RNA phylogenetics. An analysis is provided here on the *Chlamydomonas reinhardii* chloroplast 5 S RNA.

Correspondence address: N. Delihas, Department of Microbiology, School of Medicine, SUNY, Stony Brook, NY 11794, USA

## 2. THE *CHLAMYDOMONAS REINHARDII* CHLOROPLAST 5 S RNA

A comparison of the 5 S ribosomal RNAs from the cytosol of green plants and green algae shows a close sequence homology to the cytosol 5 S RNA of *Chlamydomonas* species and it has been proposed that a *Chlamydomonas*-like organism was the ancestor to the green plants [2,3]. With respect to the origin of the green plant chloroplasts, a

*Published by Elsevier Science Publishers B.V. (Biomedical Division)*
00145793/87/$3.50 

comparison of macromolecular structures of the *C. reinhardii* chloroplast to structures of the green plant chloroplasts is important in order to assess the possible origins of the plant chloroplasts. Did the higher plant chloroplasts originate from a chloroplast-containing *Chlamydomonas*-like organism or via a separate endosymbiotic event involving an intermediate cell type in the transition from protists to the higher plants? The recently published *C. reinhardii* chloroplast 5 S RNA sequence [4] reveals some highly unusual properties for a 5 S RNA. In addition, its overall primary structure and its secondary structure diverge markedly from those of the 5 S RNAs of the green plant chloroplasts. This implies that the *C. reinhardii* chloroplast and the green plant chloroplasts are not closely related. However, a comparative analysis of individual segments of the RNA, as well as group-specific signatures, suggests that rearrangements within the 5 S structural gene may have taken place. Consequently, the *C. reinhardii* chloroplast and the green plant chloroplasts could be closely related.

## 3. CHLOROPLAST AND CYANOBACTERIAL 5 S RNAs

The seven known green plant chloroplast 5 S RNA sequences [5] are highly homologous (>90% identity between most species) and some sequences are identical, e.g. those from spinach and tobacco. The chloroplast 5 S RNAs also share several unique group-specific signatures [6]. Thus, the higher plant chloroplast 5 S RNA genes are phylogenetically very stable. Cyanobacterial 5 S RNA sequences are known for four species [5]. These RNAs have high sequence homologies to the green plant chloroplast 5 S RNAs and have most of the

Table 1

Comparisons of sequence homologies

| | % homology in 5 S RNA segment | | | | | | |
|---|---|---|---|---|---|---|---|
| | 1–58 | 21–58 | 59–77 | 78–100 | 101–110 | 111–121 | 59–121 |
| *Chlamydomonas reinhardii* and other species | | | | | | | |
| Chloroplast / chloroplast<br>*C. reinhardii* / *M. polymorpha* | 83 | 92 | 21 | 79 | 36 | 64 | 47 |
| Chloroplast / cyanobacterium<br>*C. reinhardii* / *S. lividus* III | 74 | 87 | 32 | 65 | 36 | 64 | 56 |
| Chloroplast / chloroplast<br>*C. reinhardii* / *C. paradoxa* | 81 | 87 | 32 | 78 | 34 | 83 | 57 |
| Chloroplast / chloroplast<br>*C. reinhardii* / *C. ellipsoidea* | 76 | 84 | 26 | 61 | 27 | 64 | 44 |
| Chloroplast / chloroplast<br>*C. reinhardii* / *E. gracilis* | 64 | 66 | 21 | 65 | 36 | 56 | 45 |
| Other species | | | | | | | |
| Cyanobacterium / cyanelle<br>*S. lividus* III / *C. paradoxa* | 79 | 84 | 72 | 74 | 78 | 70 | 73 |
| Cyanelle / chloroplast<br>*C. paradoxa* / *M. polymorpha* | 81 | 87 | 78 | 74 | 78 | 73 | 76 |
| Cyanobacterium / chloroplast<br>*S. lividus* III / *M. polymorpha* | 78 | 87 | 78 | 57 | 89 | 73 | 69 |
| Chloroplast / chloroplast<br>*C. ellipsoidea* / *M. polymorpha* | 71 | 83 | 56 | 71 | 75 | 67 | 65 |

group-specific signatures characteristic of the chloroplast 5 S RNAs. In addition, the 5 S RNAs of *Synechococcus lividus* strains II and II also have a deletion signature (i.e. a deletion between positions 35 and 39, *Escherichia coli* 5 S RNA numbering system) characteristic only of the green plant chloroplast 5 S RNAs [6,7]. Based on the mobility of renatured and denatured samples during polyacrylamide gel electrophoresis, 5 S RNA conformations of *S. lividus* and plant chloroplasts appear to be similar, whereas both differ from the conformation of the 5 S RNA of *E. coli* [7].

The chloroplast 5 S RNA sequences from the several photosynthetic protists, however, are dissimilar in the extent of their homology to the plant chloroplast and cyanobacterial 5 S RNA sequences. The *Euglena gracilis* chloroplast 5 S RNA sequence [5] is unusual in terms of both overall sequence homology and group-specific signatures and may have arisen by an endosymbiotic event separate from that of the green plant chloroplasts [8]. Although the *Chlorella ellipsoidea* 5 S RNA [9] has an unusual insertion at position 69 and deletion between positions 87 and 88, it has a 57–62% sequence homology and most of the group-specific signatures compared with green plant chloroplast and cyanobacterial 5 S RNAs. It thus has a degree of phylogenetic relatedness to the green plant chloroplasts. The *C. paradoxa* cyanelle 5 S RNA sequence is similar to that of the 5 S RNAs from strains of *S. lividus* and both share signatures suggesting a close phylogenetic relationship [10].

When compared to the 5 S RNAs of the green plant chloroplasts and the cyanobacterium *S. lividus*, the *Chlamydomonas* chloroplast 5 S RNA reveals a very high homology at the 5′-side (positions 1–58) but a marked divergence in sequence at its 3′-half (table 1). For example, there is a high sequence homology (92%) between the liverwort *Marchantia polymorpha* [11] and *C. reinhardii* chloroplast 5 S RNAs in the region encompassing positions 21–58. In sharp contrast, segments in the 3′-half of the RNA (positions 59–77 and 101–110) reveal a random pattern of unrelated sequences, e.g. the homology between the 5 S RNAs of *M. polymorpha* and *C. reinhardii* chloroplasts is only 21% at positions 59–77 (table 1). The marked differences in homology within given segments are not found in comparisons between other

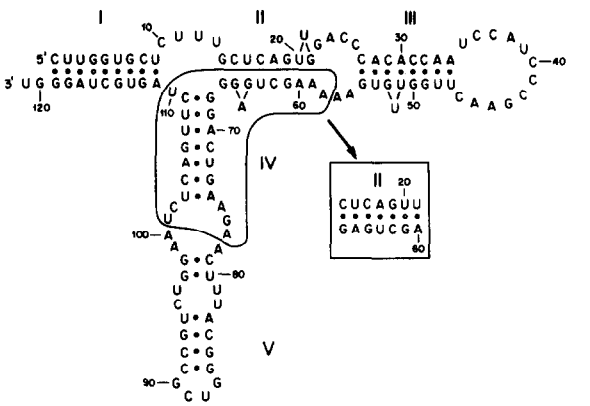


Fig.1. Nucleotide sequence [4] and secondary structure model of the *C. reinhardii* 5 S RNA. The regions encircled contain unusual sequences believed to originate from genetic rearrangements. (Inset) Possible secondary structural model for helix II [4]. The putative genetic rearrangements result in a base-pair complementarity between positions 68–74 and those at 103–109.

chloroplast/cyanobacterial 5 S RNAs (table 1).

The *C. reinhardii* chloroplast 5 S RNA sequence was fitted to the generalized 5 S RNA secondary structural model (fig.1). The 5′-side of the *C. reinhardii* chloroplast 5 S RNA has all the group-specific signatures characteristic of the chloroplast/cyanobacterial 5 S RNAs (e.g. an insertion at position A29 of the *C. reinhardii* 5 S RNA sequence and secondary structure features of helix III). The regions outside of the encircled area in the *C. reinhardii* 5 S RNA model shown in fig.1 (the 5′-half; helices I,V) have a high sequence homology to the green plant chloroplast 5 S RNAs. In contrast, particular segments of the 3′-half contain signatures that are highly unusual with respect to both the chloroplast 5 S RNAs and 5 S RNAs in general:

(i) Helix II, when drawn in the consensus secondary structural model, shows extensive mispairing of bases and a thermodynamically unstable helix (fig.1). The mispairing appears to be due to base substitutions on the 3′-side of helix II. An alternative structure for helix II [4] (fig.1, inset) is thermodynamically stable. The ability to form this type of alternate structure is generally not possible with the other

500 known 5 S RNA sequences.

(ii) The sequences encircled in fig.1, which are unusual to 5 S RNAs in general, form base-pair complementarity between positions 68–78 and 103–109. Helix IV has the potential of forming seven Watson-Crick and GoU base-pairs. There are no known eubacterial or chloroplast 5 S RNAs that can form more than four Watson-Crick and GoU pairs in this helix and most can form only two.

(iii) The universally conserved U residue of the conserved sequence $RUA_{78}$ [12] has an A substitution in the *Chlamydomonas* chloroplast 5 S RNA. With the exception of the 5 S RNA from *Sulfolobus acidocaldarius* [12] other 5 S RNA sequences do not have a substitution at this conserved U residue. Other uncommon base substitutions include those at positions 72, 74, 101 and 103.

(iv) There is an insertion between positions 107 and 110. No other known chloroplast or eubacterial 5 S RNA sequence has an insertion in this region. There is also an insertion in the region of positions 69 and 70. The unusual signatures of the *Chlamydomonas* chloroplast 5 S RNA originate from positions 59–77 and 101–110 (fig.1). Nucleotide sequences at these positions correlate with regions exhibiting a low sequence homology (table 1).

## 4. PROPOSED HYPOTHESIS FOR THE ORIGIN OF THE UNUSUAL *CHLAMYDOMONAS* CHLOROPLAST 5 S RNA

The sharp contrast in degree of identity between different regions of the *Chlamydomonas* chloroplast 5 S RNA and the 5 S RNAs of the plant chloroplasts and the cyanobacterium *S. lividus* is rarely seen in related 5 S RNAs. For example, compare the segment homologies from chloroplast, cyanobacterial, and cyanelle 5 S RNAs shown in table 1. The extreme differences in sequence homology and group-specific signatures between sections of the *C. reinhardii* 5 S RNA compared to analogous segments in other chloroplast 5 S RNAs may have taken place by genetic rearrangements within the *Chlamydomonas* 5 S gene. In an analogous situation, there exist both close similarities and marked differences between the sequences of the plant mitochondrial 5 S RNAs and the 5 S RNAs of the purple photosynthetic bacteria [14]. Here, the first 24 nucleotides from the 5′-end of the wheat mitochondrial 5 S RNA has only a 30–35% sequence homology with the equivalent segment of the 5 S RNAs of the $\alpha$ subdivision of the purple photosynthetic bacteria; the next 28 nucleotides have approx. 90% identity. A gene rearrangement in the mitochondrial genome was proposed to explain the disparity [14]. The mitochondria are believed to have originated from an ancestor to the purple photosynthetic bacteria [14–16].

Several unusual features of the ribosomal genes of *C. reinhardii* have been pointed out by Rochaix and co-workers [17]. These include an intron within the 23 S RNA gene and a fragmentation of the 5′-end of this gene into smaller 3 S and 7 S genes. The 3′-flanking side of the 5 S RNA gene was found to have unusual features of repeated sequence patterns related to sequences found in the ribosomal promoter region [4]. Thus, there are hints of major rearrangements involving the ribosomal RNA genes of *C. reinhardii* chloroplasts which may account for the unusual structure of the 5 S RNA at its 3′-side. I propose that the primary and secondary structures of segments 59–77 and 101–110 are the result of genetic rearrangements within the 5 S gene of *C. reinhardii* chloroplasts.

## 5. CONCLUSIONS

As demonstrated previously with the proposed mitochondrial origins [14], a comparative analysis of segments of the 5 S RNA can yield meaningful information not evident from a comparison of overall sequence homology. Assuming that 5 S RNA gene rearrangements took place in the *C. reinhardii* chloroplast and occurred subsequent to an endosymbiotic event, the high sequence homology and sharing of unique signatures between the 5 S RNAs of *C. reinhardii* and *S. lividus* in the 5′-half of the 5 S RNAs suggest that the *C. reinhardii* chloroplast arose by endosymbiosis involving an ancestor of a cyanobacterium related to *S. lividus*.

The 5 S RNA comparisons shown in this paper reveal that the proposal of Hori and co-workers

[2,3] for the evolution of the green plants from a progenitor organism related to *C. reinhardii* is possible if one assumes that gene rearrangements took place in *C. reinhardii* chloroplasts subsequent to a branching off of the green algae and plants.

The analysis of segments of the *C. reinhardii* 5 S RNA described here points to the importance of analyzing ribosomal RNA homologies by individual segments in phylogenetic studies and in assessing aberrant 5 S RNA secondary structures. This analysis may also be useful in higher molecular mass ribosomal RNA phylogeny and especially in eukaryotic RNAs where a large variation in G+C inserts is prevalent [18–20].

## ACKNOWLEDGEMENTS

I thank Janet Anderson and Dr James Coyer for helpful suggestions during the preparation of the manuscript. Supported by NSF grant DBM 85-02213.

## REFERENCES

[1] Delihas, N., Andersen, J. and Singhal, R.P. (1984) Progr. Nucleic Acid Res. Mol. Biol. 31, 161–190.
[2] Kumazaki, T., Hori, H. and Osawa, S. (1983) J. Mol. Evol. 19, 411–419.
[3] Hori, H. and Osawa, S. (1986) Biosystems 19, 163–172.
[4] Schneider, M., Darlix, J.-L., Erckson, J. and Rochaix, J.-D. (1985) Nucleic Acids Res. 13, 8531–8541.
[5] Erdmann, V.R. and Wolters, J. (1986) Nucleic Acids Res. 14, r1–r59.
[6] Delihas, N., Andresini, W., Andersen, J. and Berns, D. (1982) J. Mol. Biol. 162, 721–727.
[7] Delihas, N., Andersen, J. and Berns, D. (1985) J. Mol. Evol. 21, 334–337.
[8] Delihas, N. and Fox, G.E. (1987) Ann. NY Acad. Sci. 503, in press.
[9] Yamada, T. and Shimaji, M. (1986) Nucleic Acids Res. 14, 9529.
[10] Maxwell, E.S., Liu, J. and Shively, J.M. (1986) J. Mol. Evol. 23, 300–304.
[11] Yamano, Y., Ohyama, K. and Komano, T. (1984) Nucleic Acids Res. 12, 4621–4624.
[12] Delihas, N. and Andersen, J. (1982) Nucleic Acids Res. 10, 7323–7344.
[13] Stahl, D.A., Leuhrsen, K.R., Woese, C.R. and Pace, N.R. (1981) Nucleic Acids Res. 9, 6129–6137.
[14] Villanueva, E., Luehrsen, K.R., Gibson, J., Delihas, N. and Fox, G.E. (1985) J. Mol. Evol. 22, 46–52.
[15] John, P. and Whatley, F.R. (1975) Nature 254, 495–498.
[16] Yang, D., Oyaizu, H., Olsen, G.J. and Woese, C.R. (1985) Proc. Natl. Acad. Sci. USA 82, 4443–4447.
[17] Dron, M., Rahire, M. and Rochaix, J.-D. (1982) Nucleic Acids Res. 10, 7609–7620.
[18] Delihas, N. (1967) Biochemistry 6, 3356–3362.
[19] Hadjiolov, A.A., Venkov, P.V., Dolapchiev, L.B. and Genchev, D.D. (1976) Biochim. Biophys. Acta 142, 111–127.
[20] Clark, C.G., Tague, B.W., Ware, V.C. and Gerbi, S.A. (1984) Nucleic Acids Res. 12, 6197–6220.