The Possible Common Origin of tRNA and 5S rRNA

Mária Ujhelyi

Central Theoretical Department, Semmelweis University of Medicine, 1445 Budapest, Nagyvárad tér 4. P.O.B. 370, Hungary.

Z. Naturforsch. **38 c,** 501 – 504 (1983);
received January 25, 1983

tRNA, 5S rRNA, Molecular Evolution

Seryl tRNA (anticodon GCU) from mammalian mitochondria shows in comparison to other mitochondrial tRNAs additional special features differing from the generalized tRNA model. When arranged in the traditional cloverleaf form, eight bases fall within the TΨC loop, and the entire dihydrouridine loop is lacking. This seryl tRNA molecule is therefore shorter than other tRNAs. It was originally thought to represent a mitochondrial analogon of 5S rRNA and its precise classification is still disputed. The present studies suggest that this mitochondrial tRNA represents a fossil molecule which is related to the common ancestor of the present tRNA and 5S rRNA molecules.

Introduction

The transcription of tRNAs and 5S rRNAs (and of a few low molecular weight viral RNAs) is catalysed by the RNA Pol III, and has recently been shown to involve also another factor which binds specifically to an intragenic region [1, 2]. This regulation mechanism differs from the other known mechanisms of gene transcription. Similarity in the transcription mechanisms has suggested a phylogenetical relationship between tRNA and 5S rRNA molecules. The binding sites for the transcription factor have been determined and compared between tRNA and 5S rRNA to demonstrate the possible relationship, but sequencing indicated only a low homology [3]. In another examination, the flanking sequences of the Adenovirus VA I RNA control region have been identified as

5′GUGGPyNNPuGUGG…GGGUUCGAANCC3′

sequences [4]. Comparison to other RNA species revealed a high homology of these sequences to tRNAs in general, and a perfect homology to rat $tRNA^{Ser}_{GCU}$ in particular. At the same time, sequence homology to 5S rRNA was very low: the 3′ terminal sequence was partly identical, but the 5′ terminal sequence was missing altogether at the boundary of the 5S rRNA control region. The bulk of the 5′ flanking sequence falls within the D loop in tRNA. 5S rRNA thus seems to lack the segment corresponding to the D loop if the hypothesis of relationship were maintained despite negative results. However, recently certain tRNAs have been identified from some mammalian mitochondria which lack the D loop as well: as $tRNA^{Ser}_{GCU}$ [5, 6]. Originally, this molecule was thought to be a mitochondrial analogon of 5S rRNA ($3S_E$) [7]. This view was based on some similarity of $3S_E$ to mt rRNA and its relative size to the large mt ribosomal subunit. And there is an additional evidence supporting 5S rRNA equivalence: $3S_E$ in hamster mitochondria contains a region complementary to the small mt ribosomal subunit [8] as in the case of cytoplasmic 5S rRNA [9].

These ambivalent features suggest that the mt $tRNA^{Ser}_{GCU}$ might rather demonstrate the phylogenetic relationship between tRNAs and 5S rRNAs than the cytoplasmic tRNAs. It seems worthwhile to compare the primary structure of this molecule and some 5S rRNAs.

Results and Discussion

First, the sequence of mt seryl tRNA was compared to analogous cytosolic tRNAs (Table I) [10]. Secondly, the mt tRNAs were compared to the 5′ half part of 5S rRNAs from various sources. The sequences are found in Erdmann's catalogue [11]. Since the mt tRNAs differ from each other in some bases, the comparison was made one by one (Table II).

The mitochondrial $tRNA^{Ser}_{GCU}$ shows partial homology to both the analogous cytosolic tRNA and to the first 62 bases of 5S rRNAs. Bovine mt tRNA has 27 and 21 bases in common with rat cytosolic tRNA and Halobacterium tRNA, respectively. The strongest sequence homology is necessarily found in the anticodon region. In other regions sequence homology is low.

Mitochondrial seryl tRNAs and the 5′ half part of 5S rRNAs of various sources have 22 to 27 bases in common. The bulk of the homologous bases is located in the single-stranded regions in 5S rRNA secondary structure. Homology is the strongest between the TΨC and GAAC regions of tRNA and

Reprint requests to M. Ujhelyi.

0341-0382/83/0500-0501 $ 01.30/0

Table I. Nucleotide sequence of seryl tRNA from bovine mitochondria (B.mt) and that from *Halobacterium volcanii* (H.) and rat liver (R.) [10] without the CCA end. In cytoplasmic tRNA from position 10 to 26 (bases of D loop) and from 47:1 to 48 (bases of variable loop) are omitted according to the mitochondrial tRNA structure. The modification of bases is not marked. The single-stranded and double-stranded regions of the secondary structure are separated. The underlined sequences are homologous.

```
B. mt  5' GAAAAAG UAUGC AAGAA CUGCUAA UUCUA UGCU   C CCAUA UCUAAUAG UAUGG CUUUUUCG 3'
H.     5' GUUGCGG UA..G CUGGG UUGCUNA CUCAG UGGC...C CGGGG -UCGAAU- CCCCG CGGCAACG 3'
R.     5' GACGAGG UG..G AUGGA CUGCUAA UCCAU UGUG...C GUGGG UUCGAAU- CCCAU CCUCGUCG 3'
```

Table II. Nucleotide sequences of seryl tRNA from bovine (B.mt) and human (H.mt) mitochondria without the CCA end and that of rRNA from wheat mitochondria (Mt.), *Clostridium pasteurianum* (C.P.), *Anacystis nidulans* (A.N.), *Crithidia fasciculata* (C.F.), *Halobacterium cutirubrum* (H.C.), *Escherichia coli* (E.C.), *Chlamydomonas reinhardii* II (C.R.) and *Tetrahymena thermophyla* (T.T.) [11] from 1 to 60–62 nucleotide. The single-stranded and double-stranded regions of the secondary structure proposed by Fox and Woese for prokaryotic [12] and by Luehrsen and Fox for eukaryotic 5S rRNA [13] are separated. The deletions resulted from the comparison of 5S rRNAs between themselves. The underlined nucleotides indicate homology to mt tRNAs.

```
B.mt   5'    GAAAAAG UAUG-CAA GAACUG CUAAUUCU AUGC UCCCAUAUCUAAU AGUA UGGCUUUU UCG    3'
Mt.    5' AACCGGGCAC UACGGUGA GACGUG AAAACAC- CCGA UCCCAUUCCGACC UCGA UAUAUAUG UGG... 3'
C.P.   5'  UCCAGUGUC UAUGAC-- UUAGAG GUAACAC- UCCU UCCCAUUCCGAAC AGGC AGGUUAAG CUC... 3'
A.N.   5' UCCUGGUGUC UAUGGC-- GGUAUG GAACCACU CUGA ACCCAUCCCGAAC UCAG UUGUGAAA CAU... 3'
C.F.   5'  GAGUACGAC CAUACU-- UGAGUG AAAACAC- CAUA UCCCGU-CCGAUU UGUG AAGUUAAG CAC... 3'

H. mt  5'    GAGAAAG CUCA-CAA GAACUG CUAACUC- AUGC CCCCAUGUCUAAC AACA UGGCUUUC UCA    3'
H.C.   5' UUAAGGCGGC CAUAGC-- GGUGGG GUUACUC- CCGU ACCCAUCCCGAAC ACGG AAGAUAAG CCC... 3'
E.C.   5' UGCCUGGCGGC CGUAGC-- GCGGUG GUCCCAC- CUGA CCCCAUGCCGAAC UCAG AAGUGAAA CGC... 3'
C.R.   5'  AUGGAUUGCU UAUACC-- UUUAUG AAAACUC- CCCA UCCCAU-UAGCAC UGGG AAGAUAAG UCU... 3'
T.T.   5'  GUUGUCGGC CAUACU-- AAGGUG AAAACAC- CGGA UCCCAU-UCGAAC UCCG AAGUUAAG CGC... 3'
```

5S rRNA, respectively, *e.g.* 11 of the 13 bases of the GAAC loop in *E. coli* 5S rRNA are identical with the bases of TΨC arm and loop in human mt tRNA. Another highly homologous part falls within the second loop of 5S rRNAs.

On the basis of the degree of homology and distribution of homologous parts, it can be postulated that the truncated (D loop lacking) mt tRNA is related to normal tRNAs and to 5S rRNAs as well, and might thus well represent the "missing link" between the tRNAs and 5S rRNAs. Probably the common ancestor of both molecule species had been a proto-tRNA without D loop, from which the mt $tRNA^{Ser}_{GCU}$ has been conserved as a fossile molecule. The proto-tRNA could have evolved to transfer RNA and 5S rRNA by structural and functional divergence. Since their divergence from the common ancestor both molecule species have undergone innumerable mutations, thus their relationship is not obvious on direct comparison, but immediately conspicuous if mt seryl tRNA is regarded a mediator model. Without the D loop, the sequences of TΨC arm and loop get in a corresponding position to GAAC loop of 5S rRNAs and so a large homology can be discovered between them. The GAAC sequence is completely present in the TΨC loop of some tRNAs, *e.g.* valyl and methionyl tRNAs from yeast; GAAU or GAUC are found in many cases. The GUUC precedes GAAC also in several 5S rRNAs, *e.g.* Sulfolobus acidocaldarius and Tetrahymena. The characteristic initial UCCCA sequence of the 5S rRNA GAAC loop is also presented by tRNAs partly, such as $tRNA^{Arg}$ from yeast or $tRNA^{Phe}$ from chloroplast. (The sequences are listed in ref. [10, 11].)

Structural analysis of 5S rRNAs and their comparison to the truncated tRNA permit certain conclusions on their formation. The 5S rRNA molecule has the length of about 120 nucleotides, thus being twice as long as the truncated mt tRNA, which con-

Table III. Complementary base pairs in the two halves of the 5S rRNAs from wheat mitochondrion (Mt.), *Halobacterium cutirubrum* (H.C.), *Proteus vulgaris* (P.V.), *Clostridium pasteurianum* (C.P.), *Chlamydomonas reinhardii* II (C.R.) and *Crithidia fasciculata* (C.F.). The midline is always after the 13th base following upon the GAAC or the corresponding sequence, as the tRNAs also contain 13 bases after the TΨC loop.

```
Mt.  5'   AACCGGGCACUACGGUGAGACGUGAAAACAC-CCGAUCCCAUUCCGACCUCGAUAUAUAU--
          |||||| ||| ||| | |||| | | | | | ||  |  |  |  | |
     3'   AAGGCCCGAAAUGGUACAGAGGGCUUGUUAAAGUC-AUGAUAUCCGCGUUCUGCU-AAGGUG
                    C

H.C. 5'   UUAAGGCGGCCAUAGC--GGUGGGGUUACUC-CCGUACCCAUCCCGAACACGGAAGAUAAGC
          ||||||| || |    |  ||   |||  ||| || | | || |  ||  |   |
     3'   UCAUCCGCCGCUGGCCUAAAGGGUCUCCGAG-CGUGAGGUCAUGACUGGCCUUGCGUCCGCC
                   U

P.V. 5'   UGUCUGGCGGCCAUAGC--GCAGUGGUCCCAC-CUGAUCCCAUGCCGAACUCAGAAGUGAAAC
          |||||||||   |   |  | ||          ||||||||||||   || |  || |||
     3'   UACGGACCGUCAAGGGAU---GAGAGUGUACCC-CUCUGGGGUGUGGUAGUAGCCGCGAUGUUG

C.P. 5'    UCCAGUGUCUAUGAC--UUAGAGGUAACAC-UCCUUCCCAUUCCGAACAGGCAGGUUAAGC
           |||||||||| ||     ||    |||     |  |||| |  |   |  |  | ||
     3'    UGGGUCGCAGCUGGAU-GAAGAGGUGU-CCC-GAAGGGGACGUCAUGGUAGUCGUGUAAUCU

C.R. 5'   AUGGAUUGCUUAUACC--UUUAUGAAAACUC-CCCAUCCCAUU-AGCACUGGGAAGAUAAGU
          |||| ||| |  |     ||| ||      || |||||  |||   | | ||   ||  ||
     3'   UUGUCCAGCAGUGACUCCCAAGCGUGCACCAGGGGGUGGCAUGAUGACUAAGUCGGGUAAGUA

C.F. 5'    GAGUACGACCAUACU--UGAGUGAAAACAC-CAUAUCCCG-UCCGAAUUGUGAAGUUAAGC
           ||||||| |  |    |  |||     | |  ||| || | | |||||  |   |  |
     3'    CCCUCAUGCCGUGAGUCCCAAGGGCUCAGUAGUGACUGGAGUCAUGAUUGAUUCGGACACCCA
```

sists without the CCA end of 59, 60 and 60 nucleotides in the mitochondria of humans, hamsters and bovines, respectively. With the 5S rRNA molecule coiled at about position 58–60, an 50% or even greater complementarity of the two halves can be observed; for example the bisected Proteus vulgaris 5S rRNA presents 35 complementary base pairs. The complementarity is still more conspicuous, if series of 5S rRNAs are compared (Table III). In this light, the 5S rRNA may have arisen by inverse duplication of a primitive molecule closely related to the truncated mt tRNA.

This idea could explain certain particularities of 5S rRNA, among other things, intra-molecular conformational changes [9] based on the complementarity between sequence 33–42 and 79–88 numbering in *E. coli* 5S rRNA. In view of the location of the complementary sequences, their complementarity may well have been the result of inverse duplication. The latter could also account for the position (to the TΨC in tRNA) of the 3′ flanking sequence of the transcription control region of 5S rRNA between positions 71 and 80 as demonstrated by Sakonju *et al.* [1, 2].

The close structural resemblance of the truncated mt seryl tRNA to the supposed common ancestor of tRNA and 5S rRNA molecules offers an explanation for its ambivalent features which hampered its direct identification as 5S rRNA analogon or tRNA.

The present findings agree with Küntzel and Köchel's hypothesis on the independent origin of fungal and animal mitochondria based upon the differences in gene organization, tRNA structure and codon usage [14]. Evidence of conservation of a fossil molecule from the primitive ancestor of tRNA and 5S rRNA in mammalian mitochondria suggests that they may have diverged earlier from the bacterial precursor than fungal mitochondria.

Acknowledgement

The author is indebted to Prof. Dr. G. Csaba for encouragement and valuable advice.

[1] S. Sakonju, D. F. Bogenhagen, and D. D. Brown, Cell **19,** 13–25 (1980).
[2] D. F. Bogenhagen, S. Sakonju, and D. D. Brown, Cell **19,** 27–35 (1980).
[3] W. F. Doolittle, FEBS Letters **126,** 147–149 (1981).
[4] D. M. Fowlkes and T. Shenk, Cell **22,** 405–413 (1980).
[5] P. Arcari and G. G. Brownlee, Nucl. Acids Res. **8,** 5207–5212 (1980).
[6] M. H. L. de Bruijn, P. H. Schreier, I. C. Eperon, B. G. Barell, E. Y. Chen, P. W. Armstrong, J. F. H. Wong, and B. A. Roe, Nucl. Acids Res. **8,** 5213–5222 (1980).
[7] D. T. Dubin, T. H. Jones, and G. R. Cleaves, Biochem. Biophys. Res. Comm. **56,** 401–406 (1974).
[8] R. J. Baer and D. T. Dubin, Nucl. Acids Res. **8,** 3603–3610 (1980).
[9] R. Piechoki and Th. Börner, Biol. Zentralblatt **100,** 129–153 (1981).
[10] M. Sprinzl and D. H. Gauss, Nucl. Acids Res. **10,** r 1–55 (1982).
[11] V. A. Erdmann, Nucl. Acids Res. **10,** r 93–115 (1982).
[12] G. E. Fox and C. R. Woese, J. Mol. Evol. **6,** 61–76 (1975).
[13] K. R. Luehrsen and G. E. Fox, Proc. Nat. Acad. Sci. USA **78,** 2150–2154 (1981).
[14] H. Küntzel and H. G. Köchel, Nature **293,** 751–755 (1981).