Volume 12 Number 6 1984 Nucleic Acids Research

Phylogeny of the conserved 3' terminal structure of the RNA of small ribosomal subunits

P.H.Van Knippenberg, J.M.A.Van Kimmenade and H.A.Heus

Department of Biochemistry, State University of Leiden, P.O. Box 9505, 2300 RA Leiden, The Netherlands

Received 16 January 1984; Accepted 24 February 1984

ABSTRACT

The strongest conserved part of the RNA of small ribosomal subunits is probably located near the 3' end. This paper reviews the primary and secondary structures of some 40 sequenced 3' termini and tries to classify these structures according to common features and differences.

The regions under consideration contain at the 5' side an almost universal, supposedly single-stranded stretch of nucleotides with the sequence --AAGUCGUAACAAGGU--. This is followed by a stem-loop structure. The stem always contains 9 basepairs (including U-G pairs) and no mismatches or bulged nucleotides. The loop of the hairpin is either (m^2)GGm^2_2Am^6_2A (bacteria, chloroplasts and mitochondria) or UGm^2_2Am^6_2A (cytoplasm). The hairpin is, in most cases, followed at the 3' side by --GGAUCA--. Next to it bacteria and chloroplasts contain the so-called "Shine and Dalgarno" sequence --CCUCC--.

The stem region of the hairpin contains a conserved $\begin{matrix}A\text{-}U\\U\text{-}G\end{matrix}$ junction. The two basepairs between this junction and the loop are either of type 1 ($\begin{matrix}G\text{-}C\\G\text{-}C\end{matrix}$) or type 2 ($\begin{matrix}C\text{-}G\\C\text{-}G\end{matrix}$). Classification according to type links certain bacteria with mitochondria of yeast and plants and others with chloroplasts and with animal mitochondria.

INTRODUCTION

Recent proposals for the secondary structures of the large ribosomal RNAs are based in part on phylogenetic comparisons of sequence data (1-3). Stem-loop structures (hairpins) are considered to be confirmed if they can be maintained from one species to the other through coordinate basepair changes. Many supposedly single-stranded regions on the other hand show conservation of the primary structure.

The differences between these highly conserved RNAs are of course useful to study evolutionary relationships among species, including the possible endosymbiotic origin of eukaryotic organelles, the chloroplasts and the mitochondria (4,5).

We have been interested in the structure and function of an extremely conserved region of the small ribosomal subunit RNA ("16S RNA like RNA"), *i.e.* its very 3' end (6-15). This region features, next to single-stranded

Table 1. List of known sequences at the 3' termini of RNA of small ribosomal subunits

		I 10	II 20	III	II' 30	IV 40	V 50
1.	H.halo	AACAAGGU	AGCCGUAGG	GGXX	UCUGCGGCU	GGAUCA	CCUCCU
2.	H.volc	AAGUCGUAACAAGGU	AGCCGUAGG	GGAA	UCUGCGGCU	GGAUCA	CCUCCU
3.	M.ther	UAACAAGGU	AGCCGUAGG	GGAA	CCUGCGGCU	GGAUCA	CCUCU
4.	A.tume	GXXGUAACAAGGU	AGCCGUAGG	GGAA	CCUGCGGCU	GGAUCA	CCUCCUUUCU
5.	E.coli	AAGUCGUAACAAGGU	AACCGUAGG	GGAA	CCUGCGGUU	GGAUCA	CCUCCUUA
6.	P.vulg						
7.	R.trif	GUCGUAACAAGGU	AGCCGUAGG	GGAA	CCUGCGGCU	GGAUUA	CCUCCUCCUU
8.	Synech	AAGUCGUAACAAGGU	AGCCGUACC	GGAA	GGUGUGGCU	GGAUCA	CCUCCUUA
9.	B.brev	GGU	AUCCGUACC	GGAA	GGUGCGGAU	GGAUCA	CCUCCUUUCU
10.	B.stea	GUCGUAACAAGGU	AGCCGUACC	GGAA	GGUGCGGCU	GGAUCA	CCUCCUUUCUA
11.	B.subt	AAGUCGUAACAAGGU	AGCCGUAUC	GGAA	GGUGCGGCU	GGAUCA	CCUCCUUUCUA
12.	M.capr	AGGU	AUCCGUACG	GGAA	CGUGCGGAU	GGAUCA	CCUCCUUUCU
13.	Eukary cy	AAGUCGUAACAAGGU	UUCCGUAGG	UGAA	CCUGCGGAA	GGAUCA	UUA
14.	D.disc cy	AAGUCGUAACAAGGU	AUCCGUAGG	UGAA	CCUGCGGAU	GGAUCA	UUUUA
15.	C.fasc cy	AAGUCGUAACAAGGU	AGCUGUAGG	UGAA	CCUGCAGCU	GGAUCA	UUUU
16.	N.cras cy	AAGUCGUAACAAGGU	AUCCGUUGG	UGAA	CCAGCGGAUGGGAUCA		UUA
17.	A.nidu mi	AAGUCGAAAUAUGGU	UCGUGUAAU	GGAA	AUUGCACGG	-GAUGA	AUUA
18.	Yeast mi	AAGUUG-AAUACAGU	UACCGUAGG	GGAA	CCUGCGGUG	GGCUUA	UAA
19.	Bovine mi	AAGUCGUAACAAGGU	AAGCAUACU	GGAA	AGUGUGCUU	GGAUAA	AU
20.	Hams. mi	AAGUCGUAACAAGGU	AAGCAUACU	GGAA	AGUGUGCUU	GGACUA	ACA
21.	Human mi	AAGUCGUAACAUGGU	AAGUGUACU	GGAA	AGUGCACUU	GGACGA	AC
22.	Mosq mi	AAGUUGUAACAUAGU	AGAUGUACU	GGAA	AGUGUAUCU	AGAAAG	A(A)
23.	Mouse mi	AAGUCGUAACAAGGU	AAGCAUACU	GGAA	AGUGUGCUU	GGAAUA	AUA
24.	Rat mi	AAGUCGUAACAAGGU	AAGCAUACU	GGAA	AGUGUGCUU	GGAAUA	AU
25.	Wheat mi	AAGUCGUAACAAGGU	AGCCGUAGG	GGAA	CCUGUGGCU	GGAUUG	AUCC
26.	C.rein ch	AAGUCGUAACAAGGU	AGGGCUACU	GGAA	GGUGGCCCU	GGCUCA	CCUCCUUC
27.	E.grac ch	AAGUCGUAACAAGGU	AGCCGUACU	GGAA	GGUGUGGCU	GGAACA	ACUCCC
28.	Tobac ch	AAGUCGUAACAAGGU	AGCCGUACU	GGAA	GGUGCGGCU	GGAUCA	CCUCCUUU
29.	Zea ma ch						

1. *Halobacterium halobium*; determined from 16S RNA (16). XX: unidentified nucleotides, probably $m^6_2Am^6_2A$.
2. *Halobacterium volcanii*; determined from gene (17). Positions 25,26: probably $m^6_2Am^6_2A$; position 7 is modified A.
3. *Methanobacterium thermoautotrophicum*; determined from 16S RNA (18). Position 23: modified G(m^2G?); positions 25,26: probably $m^6_2Am^6_2A$.
4. *Agrobacterium tumefaciens*; determined from 16S RNA (18). X: unidentified nucleotide; position 23: modified G(m^2G?); positions 25,26: probably $m^6_2Am^6_2A$.
5. *Escherichia coli*; sequence from 16S RNA (19-21) and from gene (22,23). Positions 25,26: $m^6_2Am^6_2A$ (24); position 5: mU (19); position 23: m^2G (8).
6. *Proteus vulgaris*; sequence from 16S RNA (25). Position 5: mU; position 23: m^2G; positions 25,26: $m^6_2Am^6_2A$.
7. *Rhizobium trifolii*; sequence from 16S RNA (18). Position 23: modified G(m^2G?); positions 25,26: probably $m^6_2Am^6_2A$.
8. *Synechococcus* AN PCC 6301 (*Anacystis nidulans*); sequence from 16S RNA (26); from gene (75). Positions 25,26: $m^6_2Am^6_2A$.

9. *Bacillus brevis*; sequence from gene (27). Positions 25,26: $m^6_2Am^6_2A$ (67).
10. *Bacillus stearothermophilus*; sequence from 16S RNA (28). Position 5: (m)U; position 23: $(m^2)G$; positions 25,26: $m^6_2Am^6_2A$ (lacking in kasugamycin resistant strain, ref. 13).
11. *Bacillus subtilis*; sequence from gene (29).
12. *Mycoplasma capricolum*; sequence from 16S RNA (18). Position 23: probably m^2G; positions 25,26: $m^6_2Am^6_2A$.
13. These eukaryotic cytoplasmic RNAs include:
Saccharomyces carlsbergensis; sequence of 18S RNA from position 24 till 3' end (30). Positions 25,26: $m^6_2Am^6_2A$; positions 28,29: UC(?).
Saccharomyces cerevisiae; sequence from RNA (31) and from gene (32).
Drosophila melanogaster; sequence from 18S RNA (31) and from gene (33).
Xenopus laevis; sequence from gene (34). Position 7: m^6A; positions 25,26: $m^6_2Am^6_2A$.
Hen reticulocyte; sequence from 18S RNA (35). Positions 25,26: probably $m^6_2Am^6_2A$; position 8: reported as G, maybe sequence artefact.
Mouse; sequence from sarcoma 18S RNA (35), and via reverse transcriptase from 18S RNA from liver (36). Positions 25,26: probably $m^6_2Am^6_2A$; position 8: reported as G.
Rabbit reticulocyte; sequence from 18S RNA (35,71). Positions 25,26: $m^6_2Am^6_2A$; position 8: was reported to be G (71) or U (35).
Rat liver; sequence from 18S RNA (35,37) and DNA (38). Position 7: methylated A (in *Novikoff hepatoma*, 39); position 8: reported as G (35) and as C (38); positions 25,26: $m^6_2Am^6_2A$ (37).
Bombyx mori; sequence till $m^6_2Am^6_2A$ from RNA via reverse transcriptase (36); also from gene (40); lack of G at position 24 (?) (40).
Barley embryo; sequence from 18S RNA (35). Position 8: reported as G.
Wheat embryo; sequence from 18S RNA (77) and via reverse transcriptase (36) till $m^6_2Am^6_2A$. 3' End is 90% G.
Euglena gracilis; partial sequence from 3' end of RNA is $AUCAU\Psi G_{OH}$ (58).
14. *Dictyostelium discoideum*; sequence till $m^6_2Am^6_2A$ by reverse transcriptase of 18S RNA (36); from gene (41).
15. *Crithidia fasciculata*; sequence from 18S RNA (42). A at position 7 not modified; U at position 43 possibly modified; positions 25,26: probably $m^6_2Am^6_2A$.
16. *Neurospora crassa*; sequence from 18S RNA (43). Positions 25,26: probably $m^6_2Am^6_2A$.
17. *Aspergillus nidulans* mitochondria; sequence from gene (5).
18. Yeast mitochondria; sequence from gene (44). No methylation of A's at positions 25,26 (45,46).
19. Bovine mitochondria; sequence from mitochondrial DNA (47).
20. Hamster mitochondria; sequence from mitochondrial 13S RNA (48,49). Positions 25,26: $m^6_2Am^6_2A$.
21. Human mitochondria; sequence from mitochondrial DNA (50).
22. *Aedes albopictus* (mosquito) mitochondria; sequence from ribosomal RNA (51,52). Positions 25,26: $m^6_2Am^6_2A$.
23. Mouse mitochondria; sequence from gene and RNA (53,66). Positions 25,26: $m^6_2Am^6_2A$.
24. Rat mitochondria; sequence from 12S RNA gene (54).
25. Wheat mitochondria; sequence from 18S RNA (65). Positions 25,26: $m^6_2Am^6_2A$; position 5: modified U, possibly m^3U; position 23: possibly m^2G; heterogeneity 3' end.
26. *Chlamydomonas reinhardii* chloroplasts; sequence from the gene (55).
27. *Euglena gracilis* chloroplasts; sequence from DNA (56,57) and RNA (58). Only A at position 25 is m^6_2A (58,59); position 5: mU.
28. Tobacco chloroplast; determined from gene (60).
29. *Zea mays* chloroplast; determined from gene (61,62).

parts with a conserved sequence, a hairpin of constant length containing the characteristic $m^6_2Am^6_2A$ sequence in the loop. Since at the present time we know much more about the sequence and secondary structure of this region than of any other part of ribosomal RNA it is worthwhile to review this part separately and to consider possible implications for evolutionary relationships.

REVIEW OF SEQUENCES AND SECONDARY STRUCTURES

Since our previous review (10) a number of other sequences encompassing the 3' ends of small ribosomal subunit RNAs have been established. Table 1 and Fig. 1 give an up-to-date survey of sequences and secondary structures with reference to the original literature. Arbitrarily, we have chosen the position where the *E. coli* 16S RNA is cleaved by colicin (position 1,494 in the 16S RNA sequence) as nucleotide number 1. In general the strong conservation of 16S RNA ends two nucleotides to the 5' side (2). Note that in several cases the RNA sequence is deduced from the DNA sequence of the gene

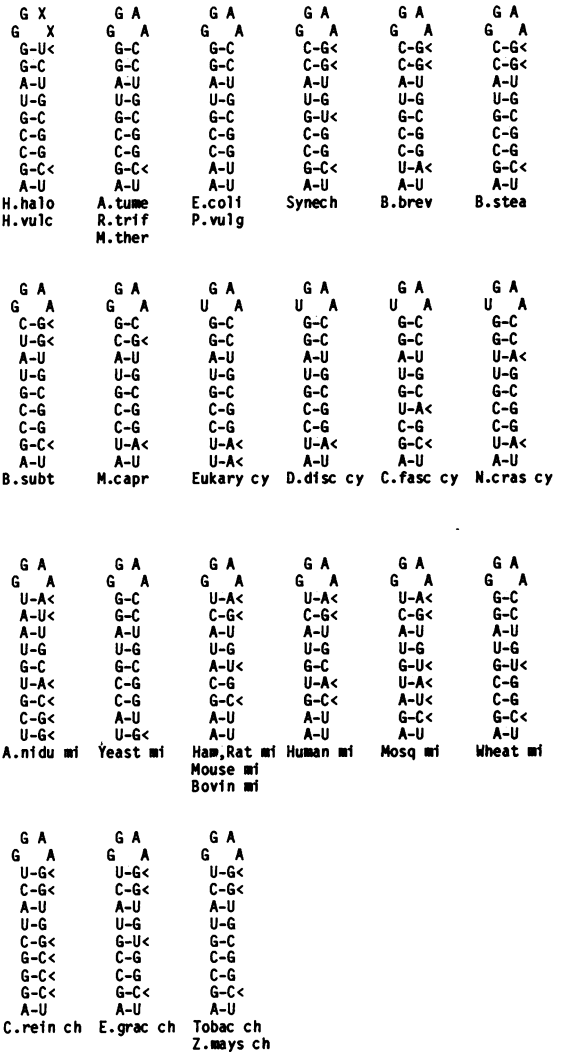


Figure 1. Stem-loop structures formed by pairing regions II and II' of Table 1. < Indicates basepair alterations in comparison with the hairpin of *E. coli*. Compare legend of Table 1 for full details. Modified nucleotides are not indicated.

and that therefore the exact 3' termini of the RNAs and the sites of modified nucleotides are not always known. Full details are given in the legend to Table 1.

In line with our previous review, we distinguish the following regions

for convenience:

I. The single-stranded region 5' to the hairpin (nucleotides -2 to 13). The "concensus" sequence here is clearly --AAGUCGUAACAAGGU--. This sequence exactly as written is probably found in all RNAs except those from fungal, human and mosquito mitochondria. The strongest deviations are found in the mitochondrial RNA of yeast having the sequence --AAGUUG-AAUACAGU--.

II. The 9 basepairs hairpin stem (nucleotides 14-22 and 27-35) will be discussed below in more detail. Note, however, that of the approximately 40 different species represented in Table 1 and Fig. 1 only Watson-Crick basepairs and U·G wobble pairs are allowed. This is probably the largest "perfect" basepaired structure in 16S RNA that is completely conserved (67).

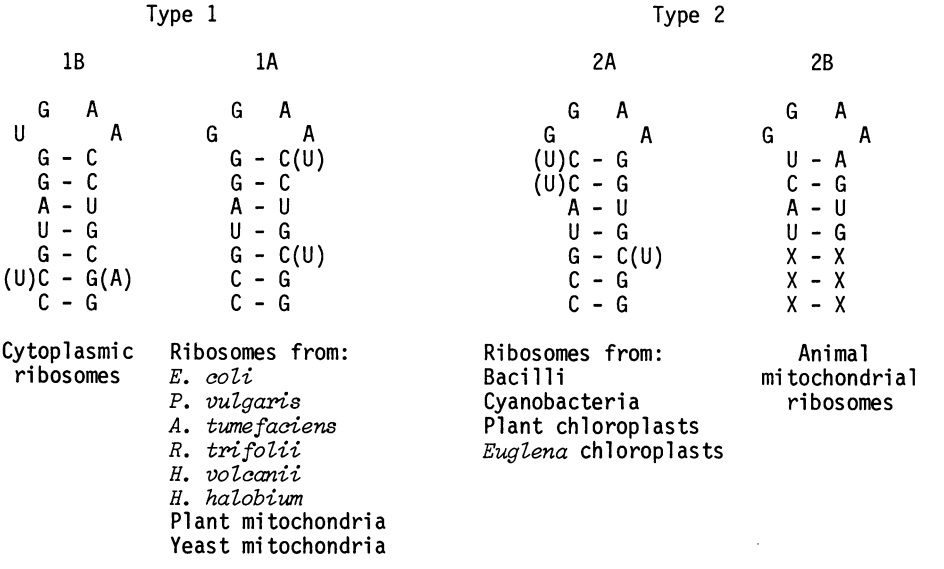
III. The four-membered loop containing either $(m^2)GGm^6_2Am^6_2A$ (bacteria, mitochondria, chloroplasts) or $UGm^6_2Am^6_2A$ (eukaryotic cytoplasm). A few exceptions are known where the A's are (partially) non-methylated: yeast mitochondrial 15S RNA contains two unmethylated A's in the loop (45,46); in chloroplasts of *Euglena gracilis* only the A at position 25 is methylated (58,59). Kasugamycin resistant mutants of *E. coli* (64) and *B. stearothermophilus* (13), also lack the methylation of the A's. N^2-monomethylation of G23 has only been established in *E. coli* (8), *P. vulgaris* (25) and *B. stearothermophilus* (28) and is likely in other bacteria (18).

IV. The region 3' to the hairpin which usually is --GGAUCA--. Deviations are mostly found among mitochondrial RNAs, but also the chloroplasts of *Chlamydomonas reinhardii* and *Euglena gracilis* do not conform to this sequence. Typically in all bacterial RNAs the above sequence is followed by the "Shine and Dalgarno" nucleotides --CCUCC--. Chloroplasts also have this sequence or a very similar one.

PHYLOGENY OF THE STEM-LOOP STRUCTURE

The hairpin stems shown in Fig. 1, with one exception (*N. crassa* cytoplasm), all have a $\begin{matrix}A\text{-}U\\U\text{-}G\end{matrix}$ junction at positions $\begin{matrix}20\text{-}29\\19\text{-}30\end{matrix}$. Proton magnetic resonance studies with the *E. coli* colicin fragment have shown that this represents a "weak spot" in the helix (14,15). The nucleotides between these conserved basepairs and the loop, with two exceptions (*Mycoplasma capricolum* and *Aspergillus nidulans* mitochondria) are either $\begin{matrix}G\text{-}C(U)\\G\text{-}C\end{matrix}$ or $\begin{matrix}Py\cdot Pu\\(U)C\cdot G\end{matrix}$. For the sake of convenience we refer to these as type 1 and type 2, respectively. All the

eukaryotic cytoplasmic hairpins are of type 1. The bacterial species *E. coli*, *P. vulgaris*, *A. tumefaciens*, *R. trifolii*, *H. volcanii*, *H. halobium* and plant and yeast mitochondria have also hairpins of type 1. On the other hand bacilli and the cyanobacteria are of type 2. Chloroplasts ($\begin{smallmatrix}U\text{-}G\\C\text{-}G\end{smallmatrix}$) and animal mitochondria ($\begin{smallmatrix}U\text{-}A\\C\text{-}G\end{smallmatrix}$) also belong to this type. That mammalian mitochondrial RNA can be grouped in one type (type 2) is remarkable indeed in view of the variations in the remainder of the stem ("below" the conserved $\begin{smallmatrix}A\text{-}U\\U\text{-}G\end{smallmatrix}$). The two RNA species that do not belong to either type (Fig. 1) are *M. capricolum* ($\begin{smallmatrix}G\text{-}C\\C\text{-}G\end{smallmatrix}$) and *A. nidulans* mitochondria ($\begin{smallmatrix}U\text{-}A\\A\text{-}U\end{smallmatrix}$). Disregarding these for the moment we may arrive at a further classification as follows:



Although we realize that we are dealing with only a very small region of a huge molecule and that the number of species is still rather limited (altogether the survey in Fig. 1 encompasses 38 species) we would like to point out that this subdivision complies with certain other observations regarding the origins of nucleus, mitochondrion and chloroplast (compare ref. 4 for a recent review):

a. Ribosomes of nuclear origin (especially in animal cells) show a far greater conservation than their mitochondrial counterparts (4,63). The universal U in the loop of the hairpin of cytoplasmic ribosomes distinguishes these from all the others. It has been known for a long time that cytoplasmic translation mechanisms differ in a number of respects from bac-

terial and organellar translation (63). The initiator tRNA of the cytoplasm *e.g.* is also different from bacterial and organellar initiator tRNA (76) and since there is evidence for the involvement of the 3' end of small subunit RNA in initiation (68), the U *vs.* G difference in the loop may be related to these differences.

b. There is compelling evidence that relates chloroplasts to cyanobacteria (summarized in ref. 4). We notice that the hairpin of *Euglena gracilis* chloroplasts differs in only one nucleotide from that of *Synechococcus* (*Anacystis nidulans*), a cyanobacterium.

c. Evolutionary tree analysis of ribosomal RNA gene sequences suggests an independent origin of fungal and animal mitochondria (5). Our evidence also points in the direction of a different bacterial origin.

d. Plant mitochondria resemble bacteria in having a 5S RNA molecule (69) and in our case their hairpin shows clear distinctions from animal mitochondria and a resemblance to fungal mitochondria. However, fungal mitochondrial ribosomes lack a 5S RNA molecule. On the other hand an initiator tRNA of wheat mitochondria resembles both eubacterial and fungal mitochondrial $tRNA_f^{Met}$ (70).

On the basis of the hairpin structure alone one might suggest the following provocative scheme (Fig. 2).

The "urkaryote" (*i.e.* the precursor of the eukaryotic nucleus) arose from a relative of present-day eubacteria (of the $\frac{G\text{-}C}{G\text{-}C}$ type) by a single nucleotide change (G → U in the loop). This ancestor has not survived to the present day (?). Endosymbiosis of certain bacteria (of the $\frac{G\text{-}C}{G\text{-}C}$ type)

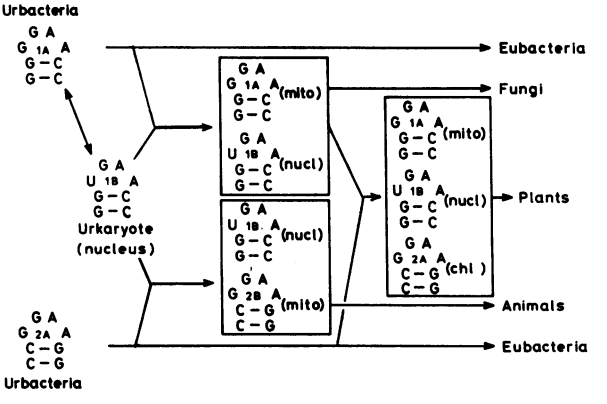


Figure 2. Possible evolutionary links between bacteria, eukaryotic cytoplasms, mitochondria and chloroplasts. Compare text.

could then have led to the evolution of the fungi and endosymbiosis of other bacteria (of the $\frac{C-G}{C-G}$ type) could have given rise to the precursor of present-day animal cells. In both cases endosymbiosis would lead to mitochondria. The development of plant cells would require a second event of endosymbiosis of $\frac{C-G}{C-G}$ type bacteria (probably cyanobacteria) into existing fungal-like eukaryotes. The more recent acquisition of chloroplasts has been invoked by some authors (72,73) to explain the stronger resemblance between bacteria and chloroplasts than between bacteria and mitochondria.
This scheme does not include the recently proposed class of archaebacteria (74). The species of which we have data (nrs. 1, 2 and 3; Table 1) show no deviation from the general eubacterial ($\frac{G-C}{G-C}$ type) sequence and they also contain the prokaryotic "Shine and Dalgarno" sequence. Since suggestions have been made that certain archaebacteria may have given rise to the nuclear lineage (4), it would be most interesting to see whether any of the archaebacteria has the typical eukaryotic U residue in the loop.

ACKNOWLEDGEMENT

We would like to thank Dr. L. Van Vloten-Doting for critical reading of the manuscript and discussions.

REFERENCES

1. Noller, H.F. and Woese, C.R. (1981) Science 212, 403-411.
2. Noller, H.F. (1983) Ann. Rev. Biochemistry (in press).
3. Stiegler, P., Carbon, P., Ebel, J.-P. and Ehresmann, C. (1981) Eur. J. Biochem. 120, 487-495.
4. Gray, M.W. and Doolittle, W.F. (1982) Microbiol. Revs. 46, 1-46.
5. Küntzel, H. and Köchel, H.G. (1981) Nature 293, 751-755.
6. Baan, R.A., Hilbers, C.W., Van Charldorp, R., Van Leerdam, E., Van Knippenberg, P.H. and Bosch, L. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 1028-1031.
7. Van Charldorp, R., Heus, H.A. and Van Knippenberg, P.H. (1981) Nucleic Acids Res. 9, 267-275.
8. Van Charldorp, R., Heus, H.A. and Van Knippenberg, P.H. (1981) Nucleic Acids Res. 9, 2717-2725.
9. Van Charldorp, R., Heus, H.A., Van Knippenberg, P.H., Joordens, J., De Bruin, S. and Hilbers, C.W. (1981) Nucleic Acids Res. 9, 4413-4422.
10. Van Charldorp, R. and Van Knippenberg, P.H. (1982) Nucleic Acids Res. 10, 1149-1158.
11. Van Charldorp, R., Verhoeven, J.J., Van Knippenberg, P.H., Haasnoot, C.A.G. and Hilbers, C.W. (1982) Nucleic Acids Res. 10, 4237-4245.
12. Heus, H.A., Van Kimmenade, J.M.A., Van Knippenberg, P.H. and Hinz, H.-J. (1983) Nucleic Acids Res. 11, 203-210.
13. Van Buul, C.P.J.J., Damm, J.B.L. and Van Knippenberg, P.H. (1983) Mol. Gen. Genet. 189, 475-478.
14. Heus, H.A., Van Kimmenade, J.M.A., Van Knippenberg, P.H., Haasnoot, C.A.G., De Bruin, S.H. and Hilbers, C.W. (1983) J. Mol. Biol. 170, 939-956.

15. Van Knippenberg, P.H. and Heus, H.A. (1983) J. of Biol. Structure and Dynamics 1, 371-381.
16. Kagramanova, V.K., Mankin, A.S., Baratova, L.A. and Bonatova, A.A. (1982) FEBS Lett. 144, 177-180.
17. Gupta, R., Lanter, J.M. and Woese, C.R. (1983) Science 221, 656-659.
18. Van Kimmenade, J.M.A., Heus, H.A. and Van Knippenberg, P.H., unpublished results.
19. Ehresmann, C., Stiegler, P., Mackie, G.A., Zimmermann, R.A., Ebel, J.-P. and Fellner, P. (1975) Nucleic Acids Res. 2, 265-278.
20. Carbon, P., Ehresmann, C., Ehresmann, B. and Ebel, J.-P. (1978) FEBS Lett. 94, 152-156.
21. Carbon, P., Ehresmann, C., Ehresmann, B. and Ebel, J.-P. (1979) Eur. J. Biochem. 100, 399-410.
22. Young, R.A. and Steitz, J.A. (1978) Proc. Natl. Acad. Sci. U.S.A. 75, 3593-3597.
23. Brosius, J., Palmer, M.L., Kennedy, P.J. and Noller, H.F. (1978) Proc. Natl. Acad. Sci. U.S.A. 75, 4801-4805.
24. Ehresmann, C., Fellner, P. and Ebel, J.-P. (1971) FEBS Lett. 13, 325-328.
25. Carbon, P., Ebel, J.-P. and Ehresmann, C. (1981) Nucleic Acids Res. 9, 2325-2333.
26. Borbély, G. and Simoncsits, A. (1981) Biochem. Biophys. Res. Common. 101, 846-852.
27. Noller, H.F., personal communication.
28. Van Charldorp, R., Van Kimmenade, J.M.A. and Van Knippenberg, P.H. (1981) Nucleic Acids Res. 9, 4909-4917.
29. Loughney, K., Lund, E. and Dahlberg, J.E. (1982) Nucleic Acids Res. 10, 1607-1624.
30. De Jonge, P., Klootwijk, J. and Planta, R.J. (1977) Nucleic Acids Res. 4, 3655-3663.
31. Shine, J. and Dalgarno, L. (1974) Biochem. J. 141, 609-615.
32. Rubtsov, P.M., Musakhanov, M.M., Zakhazyev, V.M., Krayev, A.S., Skryabin, K.G. and Bayev, A.A. (1980) Nucleic Acids Res. 8, 5779-5794.
33. Jordan, B.R., Latil-Damotte, M. and Jordan, R. (1980) FEBS Lett. 117, 227-231.
34. Salim, M. and Maden, B.E.H. (1981) Nature 291, 205-208.
35. Azad, A.A. and Deacon, N.J. (1980) Nucleic Acids Res. 8, 4365-4376.
36. Hagenbuchle, O., Santer, M., Steitz, J.A. and Mans, R.J. (1978) Cell 13, 551-563.
37. Alberty, H., Raba, M. and Gross, H.J. (1978) Nucleic Acids Res. 5, 425-434.
38. Torczynski, R., Bollon, A.P. and Fuke, M. (1981) Mol. Gen. Genet. 184, 557-559.
39. Choi, Y.C. and Busch, H. (1978) Biochemistry 17, 2551-2560.
40. Samols, D.R., Hagenbuchle, O. and Gage, L.P. (1979) Nucleic Acids Res. 7, 1109-1119.
41. Carroll, R.Mc., Olsen, G.J., Stahl, Y.B., Woese, C.R. and Sogin, M.L. (1983) Biochemistry (in press).
42. Schnare, H.N. and Gray, M.W. (1981) FEBS Lett. 128, 298-304.
43. Kelly, J.M. and Cox, R.A. (1982) Nucleic Acids Res. 10, 6733-6745.
44. Sor, F. and Fukuhara, H. (1980) C.R. Acad. Sci. Paris 291, 933-936.
45. Klootwijk, J., Klein, I. and Grivell, L.A. (1975) J. Mol. Biol. 97, 337-350.
46. Martin, N.C., personal communication.
47. Anderson, S., De Bruijn, M.H.L., Coulson, A.R., Eperon, I.C., Sanger, F. and Young, I.G. (1982) J. Mol. Biol. 156, 683-717.
48. Dubin, D.T., Baer, R.J., Davenport, L.W., Taylor, R.H. and Funko, K. (1979) in Transmethylation, Usdin, E., Borchardt, R.T. and Creveling,

C.R. Eds., pp. 389-398, Elsevier North Holland, New York.
49. Baer, R.J. and Dubin, D.T. (1980) Nucleic Acids Res. 8, 4927-4941.
50. Eperon, I.C., Anderson, S. and Nierlich, D.P. (1980) Nature 286, 460-467.
51. Dubin, D.T. (1981) in Mitochondrial genes, Attardi, G., Borst, P. and Sloninsky, P. Eds., Cold Spring Harbor Press.
52. Dubin, D.T. and Hsu Chen, C.C. (1983) Plasmid 9, 307-320.
53. Van Etten, R.A., Walberg, M.W. and Clayton, D.A. (1980) Cell 22, 157-170.
54. Kobayashi, M., Seki, T., Yaginuma, K. and Koike, K. (1981) Gene 16, 297-307.
55. Dron, M., Rahire, M. and Rochaix, J.-D. (1982) Nucleic Acids Res. 10, 7609-7620.
56. Orozco, E.M. Jr., Rushlow, K.E., Dodd, J.R. and Hallick, R.B. (1980) J. Biochem. 255, 10997-11003.
57. Graf, L., Kössel, H. and Stutz, E. (1980) Nature 286, 908-910.
58. Steege, D.A., Graves, M.C. and Spremulli, L.L. (1982) J. Biol. Chem. 257, 10430-10439.
59. Van Buul, C.P.J.J., unpublished result.
60. Tohdoh, N. and Sugiura, M. (1982) Gene 17, 213-218.
61. Schwarz, Zs. and Kössel, H. (1979) Nature 279, 520-522.
62. Schwarz, Zs. and Kössel, H. (1980) Nature 283, 739-742.
63. Kozak, M. (1983) Microbiological Revs. 47, 1-45.
64. Helser, T.L., Davies, J.E. and Dahlberg, J.E. (1971) Nature New Biol. 233, 12-14.
65. Schnare, M.N. and Gray, M.W. (1982) Nucleic Acids Res. 10, 3921-3932.
66. Van Etten, R.A., Bird, J.W. and Clayton, D.A. (1983) J. Biol. Chem. 258, 10104-10110.
67. Woese, C.R., Gutell, R., Gupta, R. and Noller, H.F. (1983) Microbiol. Revs. (in press).
68. Steitz, J.A. (1979) in Biological Regulation and Development, Goldberger, R.F. Ed., pp. 349-399, Plenum Publ. Corp., New York.
69. Leaver, C.J. and Harmey, M.A. (1976) Biochem. J. 157, 275-277.
70. Gray, M.W. and Spencer, D.F. (1983) FEBS Lett. 161, 323-327.
71. Lockard, R.E., Connaughton, J.F. and Kumar, A. (1982) Nucleic Acids Res. 10, 3445-3457.
72. Margulis, L. (1970) Origin of Eukaryotic Cells, Yale University Press, New Haven.
73. Woese, C.R. (1977) J. Mol. Evol. 10, 93-96.
74. Fox, G.E., Stackebrandt, E., Hespel, R.B., Gibson, J., Maniloff, J., Dyer, T.A., Wolfe, R.S., Balch, W.E., Tanner, R.S., Magrum, L.J., Zablen, L.B., Blackmore, R., Gupta, R., Bonen, L., Lewis, B.J., Stahl, D.A., Lulhrsen, K.R., Chen, K.N. and Woese, C.R. (1980) Science 209, 457-463.
75. Tomioka, N. and Sugiura, M. (1983) Mol. Gen. Genet. 191, 46-50.
76. Gauss, D.H. and Sprinzl, M. (1983) Nucleic Acids Res. 11, r1-r52.
77. Darzynkiewicz, E., Nakashima, U. and Shatkin, A.J. (1980) J. Biol. Chem. 255, 4973-4975.