Volume 13 Number 11 1985 Nucleic Acids Research

Further characterization of the extremely small mitochondrial ribosomal RNAs from trypanosomes: a detailed comparison of the 9S and 12S RNAs from *Crithidia fasciculata* and *Trypanosoma brucei* with rRNAs from other organisms

Paul Sloof*, Janny Van den Burg, Arthur Voogd, Rob Benne, Manuela Agostinelli[1], Piet Borst[2], Robin Gutell[3] and Harry Noller[3]

Section for Molecular Biology, Laboratory of Biochemistry, University of Amsterdam, AMC Meibergdreef 15, 1105 AZ Amsterdam, The Netherlands, and [3]Thimann Laboratories, University of California, Santa Cruz, CA 95064, USA

Received 12 March 1985; Revised and Accepted 7 May 1985

ABSTRACT

We have determined the nucleotide sequence of a maxi-circle segment from the insect trypanosome Crithidia fasciculata mitochondrial DNA, on which the genes for the major maxicircle transcripts of 9S and 12S are localized. The 5'-terminal sequences of these RNAs were determined by wandering spot analysis. The map coordinates of the 9S and 12S RNAs from Trypanosoma brucei were adjusted with respect to a previous report with the aid of primer extension analysis with reverse transcriptase. This approach allowed us to align the corresponding genes from both organisms which show an overall sequence homology of 77%. The 9S and 12S RNA genes from the two trypanosome species contain sequences, closely related to some of the regions that are universally conserved among ribosomal RNAs from members of the three primary kingdoms and their organelles, even though the overall level of sequence homology is extremely low. These universal sequences occur at positions in the 9S and 12S RNAs that are analogous to those occupied by their counterparts in authentic ribosomal RNAs. The characteristic secondary structure elements flanking these universal sequences in genuine ribosomal RNAs can also be formed in the trypanosomal 9S and 12S RNAs. These results provide unequivocal evidence for a ribosomal function of the 9S and 12S RNAs of trypanosomal mitochondria, notwithstanding their extremely small size (estimated to be 612 and 1141 nucleotides in C. fasciculata, 611 and 1150 nucleotides in T. brucei) and their unusual base composition (83% A+U).

INTRODUCTION

The biologically significant coding capacity of kinetoplast DNA (kDNA), the extraordinary mitochondrial DNA (mtDNA) of haemoflagellate protozoa (1-3), resides in the maxi-circle component (4-10). The nucleotide sequence has been determined of a major part of the Trypanosoma brucei maxi-circle. This part contains a set of genes for typical mtDNA-encoded proteins, c.f. apocytochrome b (11a,11b), the cytochrome c oxidase subunits I and II (12), unassigned reading frames (11,12), and the genes for the major maxi-circle transcripts: the 9S and 12S RNAs (13). These high abundance RNAs occur in equimolar amounts and are not specifically retained on oligo(dT) cellulose (4), properties which are

shared with mitochondrial ribosomal (r)RNAs from other organisms. A ribosomal function for the trypanosomal 9S and 12S RNAs has therefore been suggested (2,4,14). However, direct evidence for this proposal is still lacking because all attempts to isolate a 9S and 12S RNA containing ribosomal fraction from trypanosomal mitochondria have sofar been unsuccessful (4,15).

A powerful approach to investigate a possible ribosomal function for the trypanosomal 9S and 12S RNAs is to exploit the extremely high structural conservation of ribosomal RNAs (rRNAs) throughout evolution (16,17). In a previous report from this laboratory (13) this approach was followed for the 9S and 12S RNAs of T. brucei leading to the proposal that these RNAs are the trypanosomal equivalents of mitochondrial ribosomal RNAs (mtrRNAs). However, this conclusion was only tentative due to the low degree of sequence homology between 9S and 12S and other rRNAs, which made it impossible to determine the precise 5'- and 3'-end points of the 9S and 12S RNA genes. We have therefore directly determined the 5'-terminal sequences of the 9S and 12S RNAs from T. brucei and extended the analysis to the insect trypanosome C. fasciculata, from which we have also cloned and sequenced the 9S and 12S RNA genes. By sequence comparison we have investigated whether putative rRNA-like sequences in 9S and 12S RNAs are conserved among trypanosomes, and whether possible secondary structures can be proven by phylogenetic criteria (18).

In this report we show that a limited number of regions in the trypanosome 9S and 12S RNAs is homologous to sequences which are highly conserved among the members of the three primary kingdoms and their organelles (18). Moreover, characteristic secondary structures flanking these conserved sequences can also be formed in trypanosome 9S and 12S RNAs. From these structural analogies we conclude that the trypanosomal 9S and 12S RNAs are indeed ribosomal RNAs.

Trypanosomes branched from the phylogenetic tree of eukaryotic organisms at an early stage (19). Since then, trypanosomal mtDNA has apparently been exposed to different evolutionary pressures compared with mtDNAs from other eukaryotes and this is emphasized both by its unique organisation into kDNA and by a low degree of sequence homology with other mtDNAs (1-3, 11-13). The conserved sequences and their structures in trypanosomal mtrRNAs may therefore represent basic features of rRNA, which are the minimal requirements for ribosomal function.

MATERIALS AND METHODS

Materials - Restriction endonucleases were from New England Biolabs or Boehringer Mannheim; DNA polymerase (large fragment), calf intestine phosphatase and T4 DNA ligase from Boehringer; Avian myeloblastosis virus reverse transcriptase from Dr.J.Beard (Life Sciences Inc., St.Petersburg, Florida, USA); exonuclease Bal-31 from New England Biolabs or Bethesda Research Laboratories; Low-melting agarose from Bethesda Research Laboratories.

DNA and RNA - Relevant data on the procedures for Crithidia fasciculata cultivation and kDNA isolation are described in ref.14 or refs therein. Total cellular RNA from Trypanosoma brucei strain 427 (culture form) and C. fasciculata were isolated with the LiCl precipitation method (20). RNA was stored at -20°C as an ethanol precipitate. Plasmid and M13RF DNAs were isolated according to ref.21.

Cloning in plasmid and bacteriophage M13 DNA

C. fasciculata kDNA was cleaved with restiction endonuclease HindIII and the resulting fragments were ligated into plasmid pBR322. From the HindIII-C. fasciculata maxi-circle bank a clone containing the D5D1 fragment as insert (pD5D1) was selected by size, restriction enzyme fragmentation pattern and hybridization with T. brucei fragment RR3, which contains the 9S and 12S RNA genes (4) (see Fig.1). This recombinant plasmid was used as starting material for the sequence determination. p(D5D1) was cut with SalI and PstI and the resulting fragments were separated by electrophoresis on a low melting agarose gel. The 3 kb fragment S2-PS1 (see Fig.1) was isolated from the gel according to ref.22. M13 clones containing a varying part of the fragment S2-PS1 were generated with Bal31 exonuclease using non-random cloning procedures according to ref.23, except that gel isolation of Bal31 treated fragments was omitted (see ref.12). The ligation reactions were performed with vectors treated with calf-intestine phosphatase according to ref.12. This approach yielded 55 M13 clones containing a series of overlapping inserts from both strands. These clones were subjected to DNA sequence analysis.

DNA sequence analysis

DNA sequence determination was carried out by the dideoxynucleotide chain-termination technique (24). In this way a number of overlapping sequences were obtained from both strands, encompassing the 9S and 12S

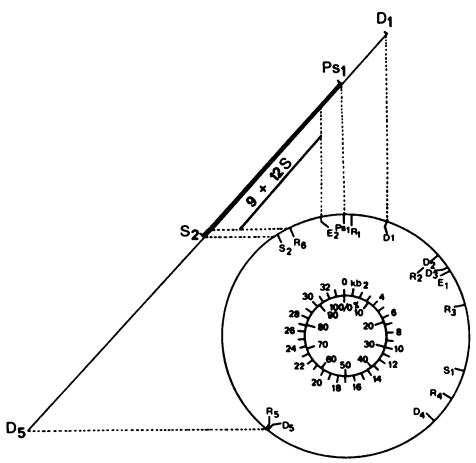

Figure 1: Physical map of the C. fasciculata maxi-circle

The 9S and 12S RNAs were previously mapped on the EcoRI-HaeIII fragment R6-E2 (15). The HindIII fragment D5 D1 was cloned in E. coli with pBR322 as vector. From this clone the SalI-PstI fragment S2-Ps1 was subcloned in bacteriophage M13. Relevant restriction sites are indicated: HindIII = D; EcoRI = R; SalI = S; HaeIII = E.

RNA genes. In regions where the sequence was ambiguous, reliable sequences were obtained by priming with specifically constructed synthetic oligonucleotides as primers for DNA synthesis (see ref.25).

5'-termini of 9S and 12S RNAs of C. fasciculata

9S and 12S RNA were separated by gelelectrophoresis and extracted from the gel. RNA was labelled with (^{32}P)-ATP using phage T4 polynucleotide kinase according to ref.26. End-labelled RNA was partially degraded by nuclease P1 and the oligonucleotides were subjected to electrophoresis on cellulose acetate followed by homochromatography on DEAE thin layer plates in the second dimension (27).

5'-termini of 9S and 12S RNAs of T. brucei

Primed cDNA synthesis using synthetic oligonucleotide as primers was carried out in the presence of dideoxy NTPs as chain terminators, essentially as described (28). The primers were located at a distance of 75-100 nucleotides from the putative 5'-end of the RNAs as determined in S1 nuclease protection experiments (29).

RESULTS AND DISCUSSION

I The nucleotide sequences of the C. fasciculata 9S and 12S RNA genes and their alignment with the T. brucei genes

The nucleotide sequence of a C. fasciculata maxi-circle segment containing 12S and 9S RNA genes, is presented in Fig.2. The 5'-ends of the corresponding RNAs were directly determined by wandering spot analysis (Fig.3). The 5'-UUUAUCA sequence for 12S RNA and the 5'-UUUUUAU sequence for 9S RNA occur only once each in the sequenced maxi-circle segment. The 5'-ends of the genes in T. brucei were determined by primer extension analysis with reverse transcriptase (see Fig.4). The 3'-ends of the 12S and 9S RNA genes of T. brucei were previously determined by S1 protection experiments at positions 1183 and 1807 respectively (29). The sequence homology in the vicinity of these positions in T. brucei and C. fasciculata maxi-circle segments drops from 100% to a background value (Fig.2). From this sequence comparison we deduced the 3'-ends of the 12S RNA genes at position 1174 in C. fasciculata and at 1181 in T. brucei, and of the 9S RNA genes at position 1834 and 1810 in C. fasciculata and T. brucei respectively. This leads to a length of the gene for the C. fasciculata 12S RNA of 1141 nucleotides (1150 nucleotides in T. brucei) and for the 9S RNA of 612 nucleotides (611 nucleotides in T. brucei), and an intergenic region of 48 nucleotides (29 nucleotides in T. brucei).

A comparison of the aligned 9S and 12S RNA gene sequences of T. brucei and C. fasciculata shows occasional small insertions and deletions (of 5 bp or less). Sequence divergence of 22% and 24% for the 9S and 12S RNA genes respectively is limited mostly to individual base replacements, which are non-randomly distributed along the genes. In the 9S RNA genes

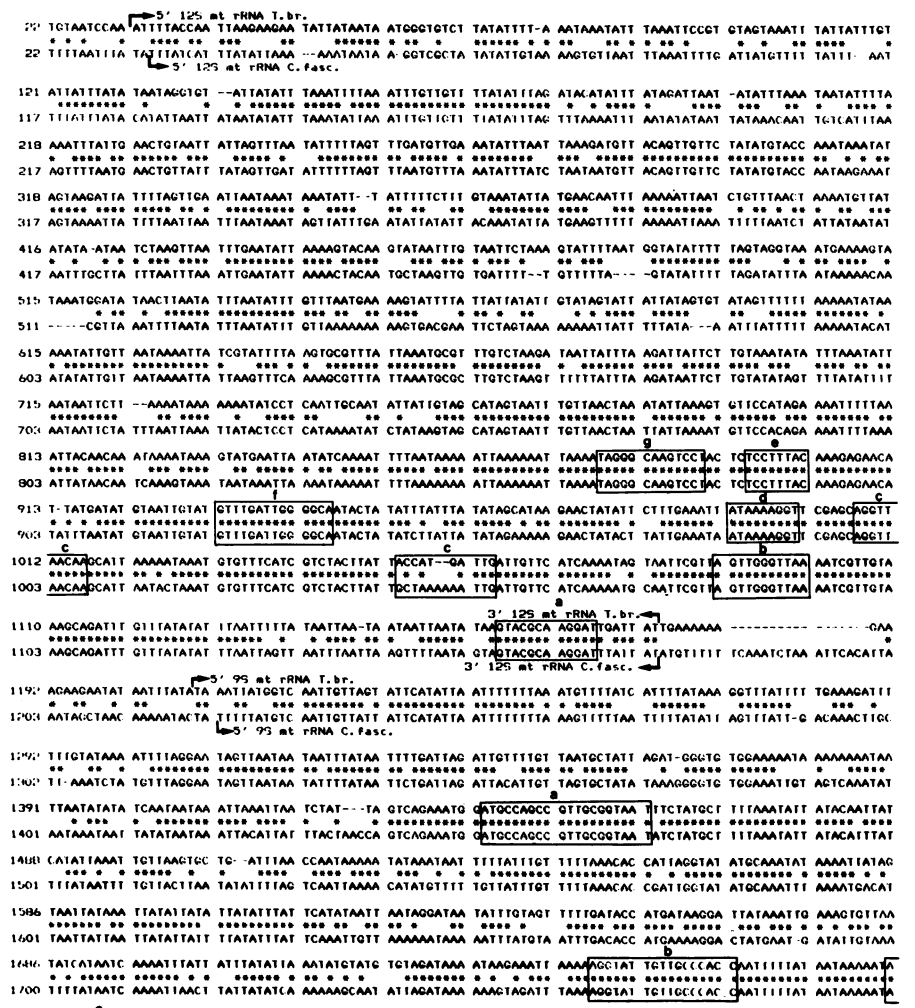

Figure 2: Aligned maxi-circle DNA segments, containing the 12S and 9S RNA genes from T. brucei and C. fasciculata.

The sequence of the T. brucei maxi-circle DNA segment containing the 9S and 12S genes (top line) is algined with the corresponding maxi-circle DNA Segment from C. fasciculata (bottom line). Identified universal sequences are boxed and indicated as in Tables II and III.

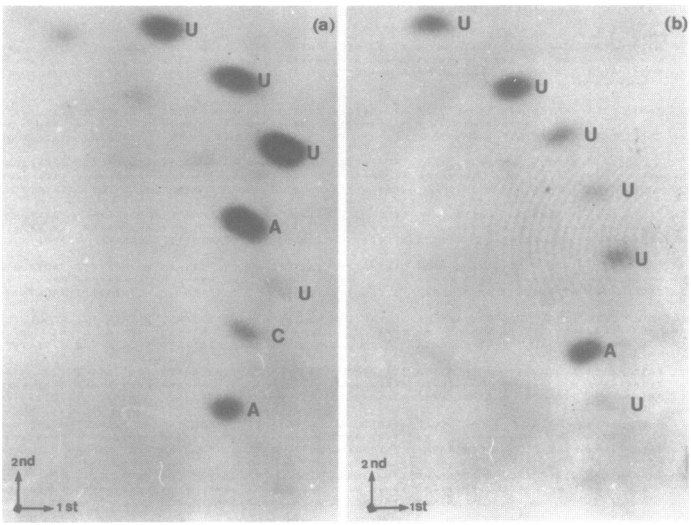

Figure 3: 5'-end determination of C. fasciculata 12S and 9S RNAs.
C. fasciculata 12S and 9S RNAs were purified electrophoretically and subjected to wandering spot analysis as described in Methods. (a) Autoradiogram of 12S RNA. (b) Autoradiogram of 9S RNA.

mutations are considerably less frequent in the regions 1427-1482 and 1747-1801 (T. brucei coordinates), whereas in the 12S RNA gene the 3'-terminal region is well conserved (T. brucei coordinates 759-1180). These highly homologous regions harbour sequences and potential secondary structures that are recognizable analogues of conserved secondary structural elements of rRNA as described in refs. 16 and 17 (see below).

Analysis of sequence divergence in protein-coding genes, tRNA- and rRNA-genes of human and bovine mtDNA, has shown that transitions appear more frequently than transversions (30). This tendency has also been observed in a comparison of the mitochondrial rRNA genes of two Paramecium subspecies (31,32). In trypanosome 9S and 12S RNA genes transversions are more prominent (15% and 16% of the bases for the 9S and 12S RNA genes respectively) than transitions (5% of the bases for both genes) (see Table I). This may be a reflection of the extreme bias towards A+T in mtDNA of trypanosomes (83% in both 9S and 12S RNA genes). The percentage sequence difference between the 9S and 12S RNA genes of the two trypanosome species (about 23%, see Table I) is as large as between the mtrRNA genes of Human and Cow (30). Assuming an equal rate of sequence evolution in mtrRNA genes of mammals and of trypanosomes we

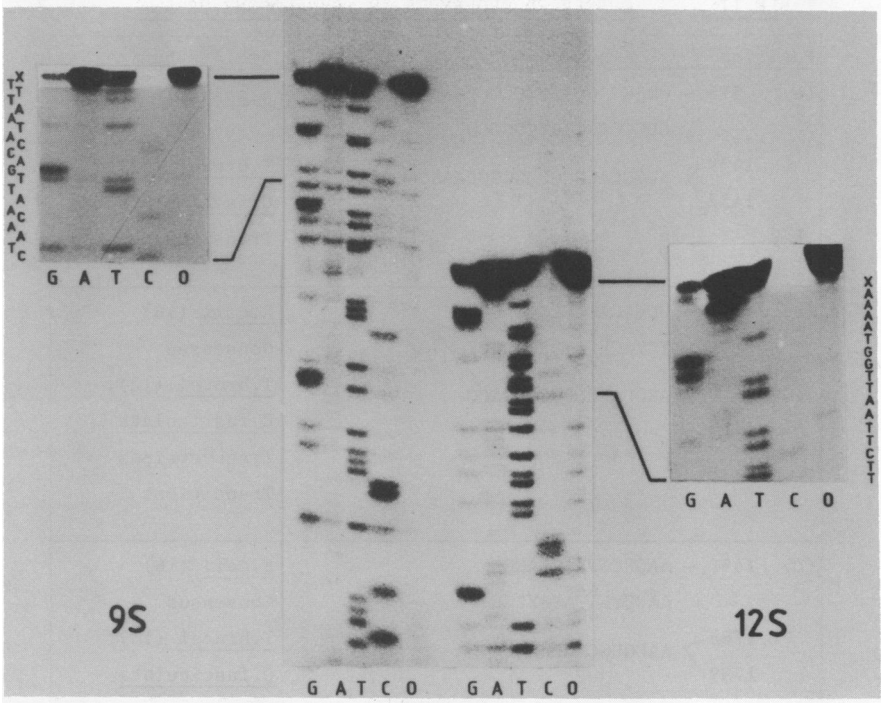

Figure 4: 5'-end determination of 9S and 12S rRNA from T. brucei.
cDNA synthesis using a synthetic dodecadeoxyribonucleotide as primer, carried out as described in Methods. Primers were:
12S RNA- 5'-TTTACTACACGG-3' (complementary to 97-108, see Fig.2).
9S RNA- 5'-ATTAACTATTCC-3' (complementary to 1308-1319, see Fig.2).

speculate that the divergence time for C. fasciculata and T. brucei is the same as for Human and Cow, namely 80×10^6 years (33).

II Structural relationship between trypanosome 9S RNA and small subunit ribosomal RNA

In order to determine structural relationships between the

TABLE I MICRO EVOLUTION OF TRYPANOSOME 9S AND 12S RNAs

	Gene	
Type of mutation	9S RNA	12S RNA
Transition	36 (5.8)	66 (5.7)
Transversion	91 (14.8)	180 (15.8)
Deletion/Insertion	9 (1.6)	32 (2.8)
Mutations total	136 (22.2)	278 (24.3)

The T. brucei and C. fasciculata mitochondrial 9S and 12S RNA genes were aligned (see Fig.2) and the observed differences tabulated according to the type of mutation. Numbers in parentheses represent the percentage of bp's of the given gene with the respective mutation.

TABLE II CONSERVED SEQUENCES IN TRYPANOSOME 9S RNA

Position	Sequence	Small subunit rRNA
(a) 515 -	GUGCCAGCAGCCGCGGUAAU	E.coli (16)
	GUGCCAGC$^{A}_{C}$RYCGCGGU$^{A}_{C}$AY	Consensus
1439 > 1452	AUGCCAGCCGUUGCGGUAAU	T.brucei (13); C.fasciculata
	* *	Transitions
(b) 1391 -	UGUACACAC^{o}GCCCGUC	E.coli (16)
	YGYAC$^{AY}_{UA}$AYYRCYCRUC	Consensus
1750 > 1764	AGGUAUUGUUGCCCACC	T.brucei (13); C.fasciculata
	* ***	Transversions
	* *	Transitions
(c) 1492 -	AAGUCGUAACAAGGU	E.coli (16)
	AAGUYG$^{AA}_{UC}$AYAXRGU	Consensus
1785 > 1799	AACGUGCAGUAAUUA	T.brucei (13); C.fasciculata
	** ***	Transversions
	* *	Transitions

Conserved sequences from trypanosome 9S RNA are shown aligned with their homologues which are universal in the 16S-like rRNAs of the three major kingdoms (typified by E. coli). Consensus sequences take into consideration the corresponding regions from mitochondrial small subunit rRNAs from Paramecium (31); Yeast (46); Aspergillus (47); Maize (48); Human (49); Cow (30); Mouse (50); Rat (51) and Mosquito (only in (b) and (c); ref.37)
R = purine; Y = pyrimidine; X = purine or pyrimidine; * indicates positions at which a transition or transversion occurred in trypanosome 9S RNA with respect to the consensus sequence; o indicates the position at which a P-site tRNA has been cross-linked to E. coli 16S rRNA (35).

trypanosomal 9S and 12S RNA and rRNAs three approaches have been followed:

(i) A search in trypanosome 9S and 12S RNAs for universal sequences found in all rRNAs (see ref.18); (ii) a search for conserved purine-pyrimidine sequences; and (iii) a search for secondary structure elements conserved in all rRNAs.

It has been shown for T. brucei that the level of sequence homology

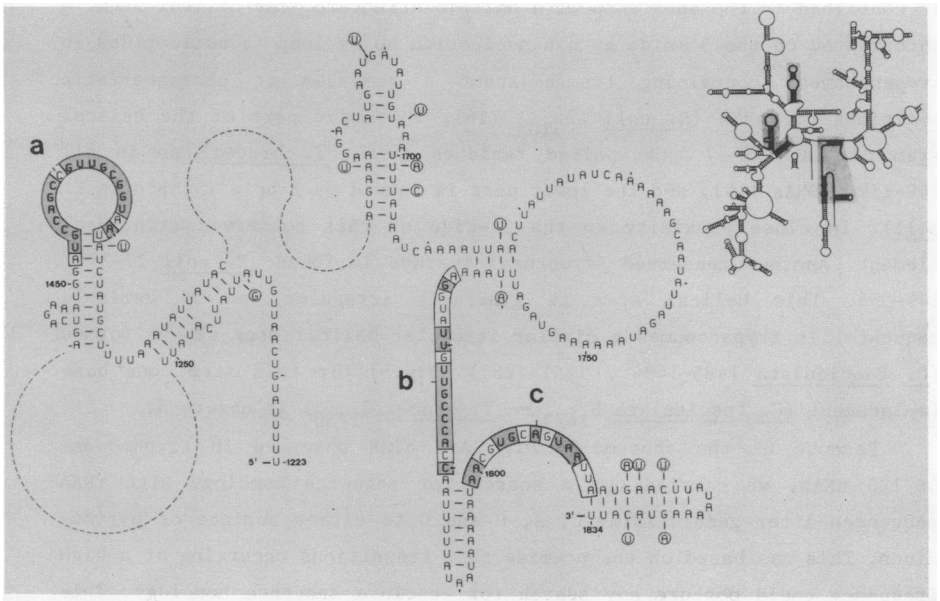

Figure 5: Conserved secondary structure features in trypanosome 9S RNA.
The sequences of relevant C. fasciculata 9S RNA regions are given. Conserved sequences a,b and c, specified in Table II, are boxed. Constant nucleotides are shaded. Differences with the corresponding T. brucei 9S RNA regions are indicated by encircled nucleotides for base replacements. A diagram of E. coli 16S rRNA is given for reference.

between the trypanosome 9S RNA and small subunit rRNA from either E. coli or mitochondria from a variety of organisms is low (13). However, one region (C. fasciculata 1452-1471) shows considerable homology to a highly conserved region of rRNA (E. coli 515-534) (16). In order to take the degree of variability in this universal sequence into account, we have derived a consensus sequence using the 515-534 region of E. coli 16S rRNA and corresponding regions in mitochondrial small subunit rRNAs from a number of organisms. A comparison of the corresponding trypanosome 9S RNA region with the consensus sequence shows only two transition mutations (see Table II, panel a).

Since the secondary structure of this region and flanking sequences (E. coli positions 500-545) is also highly conserved in the 16S-like rRNA from all organisms and organelles (16,34), we have investigated trypanosome 9S RNA in this respect. Fig.5 shows that such a secondary structure is possible. By analogy with the model for E. coli 16S rRNA (16,34), the conserved primary sequence (E. coli 515-534 / C. fasciculata 1452-1471)

is contained in the apex loop of a hairpin structure. The helical stem is interrupted on the 5'-side by a 6 nucleotide bulge loop (5 nucleotides in trypanosomes) containing two adjacent A residues at characteristic positions for rRNA (E. coli AA_{510}) (16). The upper part of the helical stem consists of 7 base paired residues (6 in T. brucei) as in all 16S-like rRNAs (16), and the lower part is formed by 3 bp's (5 bp's in E. coli). In close proximity to the 3'-side of this conserved structural element, another conserved structural feature is found: E. coli 27-37 / 547-556. This helical stem is generally irregular and of variable sequence. In trypanosomes a similar irregular helical stem can be formed (C. fasciculata 1485-1494 / 1241-1251) (Fig.5). In this stem, one base replacement (C. fasciculata U_{1473} — T. brucei G_{1230}) is observed.

Because of the abnormally high A/U bias observed in trypanosome 9S/12S RNAs, we carried out a search for sequence homology with rRNA sequences after generalizing G, A, C and U to either purines or pyrimidines. This was based on the premise that transitions occurring at a high frequency could obscure any search for absolute sequence homology. This approach led to the identification of two regions of high homology proximal to the 3'terminal part of trypanosome 9S RNA (see Table II-b and c).

The first region corresponds to positions 1391-1407 of E. coli 16S rRNA and shows 4 transversions and two transitions with respect to the latter (see Table II-b). This E. coli 16S rRNA region is located at the subunit interface (18). Furthermore, the anticodon of a P site-bound tRNA has been cross-linked to C_{1400}, indicating that codon-anticodon recognition occurs in close proximity to this position (35). A C—U transition is found at this position in trypanosome 9S RNA (see Table II), as in the mitochondrial small subunit rRNAs of two Paramecium subspecies (31). Transversions occur at the 5'-flank of this pyrimidine, while in the 3'-flanking sequence only the penultimate residue shows a transitional difference (Table II). This suggests that in E. coli the 16S rRNA sequence most crucial to the codon-anticodon recognition process includes positions located directly 3' to C_{1400}.

The second region shows strong homology to positions 1492-1503 of E. coli 16S rRNA, also located at the subunit interface (18). Trypanosome 9S RNA shows a greater divergence from the consensus here than for the other two universal sequences (Table II-c). However, as is shown below, this entire 3'-terminal region of 9S RNA can be folded in a fashion closely

resembling the model proposed for the 3'-terminal region of E. coli 16S rRNA.

Two universal sequences (E. coli 1391-1407 and 1492-1503) exist in a single stranded conformation in 16S rRNA (16) and are embedded in a characteristic set of helices (16,36). Upstream from E. coli position 1391 a long range interaction occurs (16) whereas between the two universal sequences a helical stem is found which varies in length between 7 bp (in Mosquito mitochondria, ref.37) and 49 bp (in the Rat cytoplasm, ref.38). A 3'-terminal helical stem is present in all small subunit rRNAs (16,36). This characteristic configuration can also be drawn for trypanosomal 9S RNA (Fig.5). A long-range interaction between 1762-1759 / 1714-1717 (C. fasciculata), containing one compensating base replacement relative to T. brucei at position 1716/1760, can be formed adjacent to the conserved sequence 1764-1780 (Table II). Between this conserved sequence and the conserved region 1799-1813 (C. fasciculata; Table II) a helical stem containing 7 basepairs can be drawn at precisely the predicted position. At the 3'-end of C. fasciculata 9S RNA a helical stem analogous to the 3'-terminal helix of 16S-like rRNAs can be formed which exhibits two compensating base changes relative to T. brucei 9S RNA (see Fig.5), proving their existence in vivo according to phylogenetic criteria (18). In Crithidia this helical region is interrupted, as also is observed in the mitochondrial small subunit rRNAs of two Paramecium subspecies (31). This interruption at A_{1817} / A_{1830} does not occur in T. brucei which has an A-U pair at this position. The position of the helical stem, its length (9 bp) and the number of residues in the loop (4 nucleotides) closely resemble the 3'-terminal helix found in all small subunit rRNAs (16). However, the conserved (A-U/U-G) part of the stem (39) at position: 1513-1522/1512-1523 of E. coli is not present in trypanosomes and from the conserved GGAA or UGAA sequences in the loop only the adjacent A residues are found in trypanosome 9S RNA. The adjacent A residues at this position are methylated in all rRNAs and are found unmethylated in kasugamycin resistant mutants (39). The nucleotide sequences in this conserved secondary structure element in small subunit rRNAs are related and have been used to construct a scheme of possible evolutionary links between bacteria, eukaryotic cytoplasms, mitochondria and chloroplasts (39). The nucleotide sequence of this element in trypanosome 9S RNA does not fit in this scheme of related sequences which provides evidence that trypanosomal 9S RNA has been exposed to a

(a) Small subunit rRNA	Number of nucleotides in a	b	c	Base pairs in d	e
E. coli 769-810	9	7	4	4	7
Consensus	9	6-8	3-5	3-4	6-8
T. brucei	9	6	3	3	6
C. fasciculata 1671-1705	8	6	3	3	6

(b)

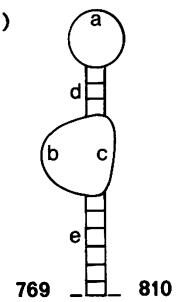

Figure 6: The conserved secondary structure of the 769-810 (E.coli) region of the central domain of small subunit rRNA

The number of nucleotides in single- or double-stranded conformation for this region in E. coli, and a consensus of the corresponding region for mitochondrial small subunit rRNA from Paramecium (32), Yeast (46), Aspergillus (47), Maize (48), Human (49), Cow (30), Mouse (50), and Rat (51) is given in (a). A schematic drawing of this region is given in (b) and regions of T. brucei and C. fasciculata 9S RNA, fulfilling the criteria formulated in the consensus are presented in (a).

different evolutionary pressure than mitochondrial small subunit rRNAs of other organisms.

From the identification of these three highly conserved small subunit rRNA sequences at analogous positions in trypanosome 9S RNA and from the potential of their flanking regions to form secondary structures closely analogous to that proposed for 16S-like rRNA (16) we conclude that 9S RNA is the trypanosome equivalent of mitochondrial small subunit rRNA. Therefore, we have searched for additional conserved secondary structure elements in trypanosome 9S RNA. A candidate is the 769-810 area (E. coli) which contains a constant secondary structure present in all rRNAs (see Fig.6). In a matrix plot of all possible base-pairing interactions within C. fasciculata 9S RNA, the only combination of single stranded and helical regions that fulfilled the criteria listed in Fig.6, is located between positions 1671-1705. In C. fasciculata the loop contains 8 residues (in T. brucei: 9 nucleotides) with the first stem being 3 bp, the second 6 bp in length, containing respectively one and

two compensating base replacements relative to T. brucei. The two helical regions are interrupted by an internal loop containing 6 nucleotides on the 5'-side and 3 nucleotides on the 3'-side, within the range of size variation observed for the consensus structure (Fig.6). The possibility of folding this region into a rRNA-like structure as well as its position relative to the other conserved rRNA sequences, suggest this region as a possible candidate for the characteristic 790 region of the central domain of small subunit rRNAs.

In comparison with mitochondrial small subunit rRNAs from other organisms the trypanosome 9S RNAs are the smallest characterized thus far. The most drastic size reductions in 9S RNA relative to E. coli 16S rRNA have occurred in the (E. coli) areas: 38-500 (58%); 810-926 (94%) 933-1383 (91%) and 1409-1491 (80%). These regions are phylogenetically the most variable; in mammalian mitochondrial small subunit rRNAs the size reductions also occur within these regions (16). Among the conserved sequences found in small subunit rRNA, the three regions identified in trypanosome 9S RNA correspond to the most highly conserved ones (16). This implies that the other conserved sequences are deleted. However, a drastic change in their sequence or the possibility that they have never been present in trypanosomes cannot be excluded.

III Structural relationship between trypanosome 12S RNA and large subunit rRNA

Some highly conserved sequences, located in the functionally important 3'-terminal region of all large subunit rRNAs (17,34) (domains V and VI of E. coli 23S rRNA, ref.17) are also found in the 3'-terminal region of trypanosome 12S RNA.

First, a 12-nucleotide sequence, which contains the α-sarcin cleavage site, implicated in the binding of the EF-1.GTP.aa-tRNA ternary complex (40,41) is present at an analogous position in trypanosome 12S RNA, with only one transition base replacement with respect to the consensus (see Table III-a). This region can be folded in a secondary structure, exposing the putative α-sarcin cleavage site in the apex loop, as has been determined for E. coli and Bacillus stearothermophilus (42) (see Fig.7). However, phylogenetic evidence for such folding in trypanosomes is lacking because of the high sequence homology between T. brucei and C. fasciculata in this area of 12S RNA.

Second, in E. coli the site on 23S rRNA which is located at the peptidyl transferase center has been identified recently (43) in a

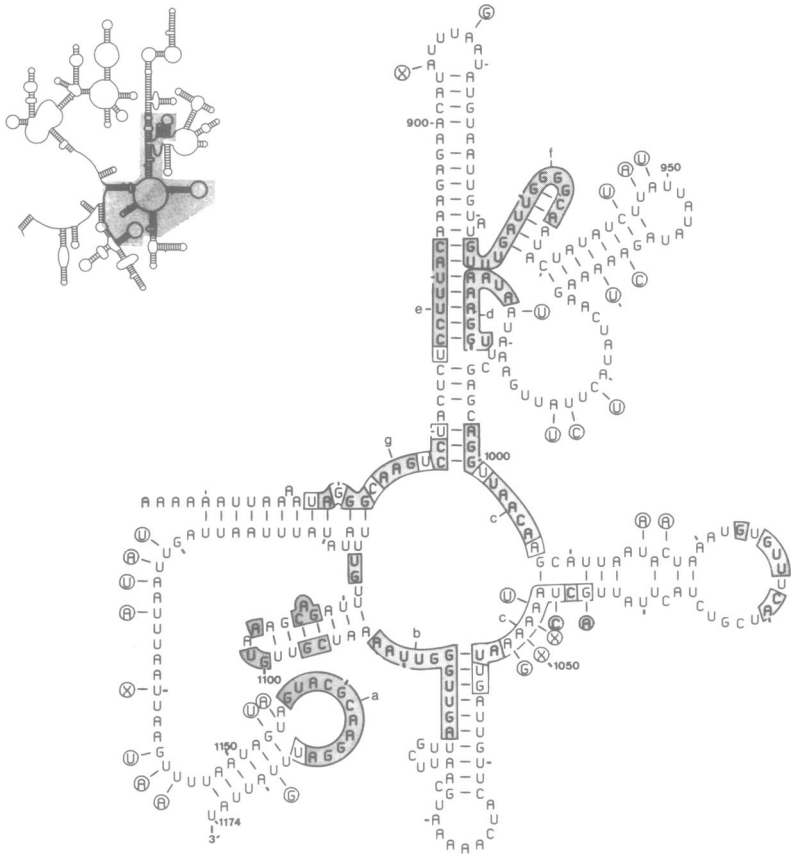

Figure 7: Secondary structure of the 3'-terminal region of trypanosome 12S RNA.

The C. fasciculata sequence is given. Conserved sequences a,b,c,d,e, f and g, specified in Table III, are boxed. Constant nucleotides are shaded. Differences with the corresponding T. brucei 12S RNA region are indicated by encircled nucleotides for base replacements and by encircled X's for deletions.

A diagram of the secondary structure of the corresponding region of 23S rRNA of E. coli (domains V and VI) is given for reference.

conserved 11 nucleotide sequence. This sequence is also present in trypanosome 12S RNA at a unique and analogous location (see Table III-b). Further support for the involvement of this region of 12S RNA in peptidyl transferase activity comes from a comparison with two rRNA sequences associated with the sensitivity of protein synthesis towards chloramphenicol (CAP), an antibiotic that inhibits peptidyl transferase activity

(44) (see Table III-c). Although these regions of trypanosome 12S RNA differ in sequence their homology with sequences in the peptidyl transferase region is nevertheless clearly apparent. Since mutants in these sequences in yeast and mammalian mitochondria confer chloramphenicol resistance, trypanosome mitoribosomes might be expected to show CAP resistance. Indeed, protein synthesis in trypanosomes appears to be resistant to CAP (29). Also the region of 23S rRNA associated with erythromycin sensitivity (45), can be identified in trypanosome 12S RNA in a slightly altered form (see Table III-g). This observation is consistent with the absence of erythromycin-sensitive protein synthesis in trypanosomes (29). Finally, three other conserved sequences in the 3'-terminal region have been identified (see Table III-d,e,f).

Using these sequences as landmarks in folding excercises on trypanosome 12S RNA we find that a secondary structure can be drawn which is similar to the 3'-terminal region of E. coli 23S rRNA (17) (see Fig.7). The putative CAP region, erythromycin region and the three other conserved regions (see Table III) can be positioned in analogous locations in the secondary structure model (Fig.7). Furthermore, a long-range RNA interaction (C. fasciculata 867-882/1130-1116) analogous to E. coli 2057-2043/2625-2611 is possible linking the putative α-sarcin region (Table III-a) and other conserved sequences (Tables III-b-g) as in the E. coli model (17). However, conclusive phylogenetic evidence for this folding in trypanosomes, is lacking because of the high sequence homology between T. brucei and C. fasciculata in this area on the one hand, and the extreme sequence divergence with respect to the large subunit rRNAs from other organisms and organelles on the other. From the presence of conserved large subunit rRNA-like sequences in the 3'-half of 12S RNA and from the possibility of folding this area in a closely homologous fashion to 23S rRNAs we conclude that 12S RNA is the trypanosomal equivalent of mitochondrial large subunit rRNA.

The trypanosome 12S RNAs are the smallest mitochondrial large subunit rRNAs characterized thusfar. The 3'-terminal region of 12S RNA is only 315 nucleotides long, whereas the equivalent region in E. coli 23S rRNA, domains V and VI, spans 862 nucleotides. The size reduction of this region in trypanosome 12S RNA relative to E. coli 23S rRNA, is not evenly distributed. Some parts are constant while others are drastically reduced in size or are eliminated. The functionally important central part of this region, in which some sequences are identified that are involved in

TABLE III CONSERVED SEQUENCES IN TRYPANOSOME 12S RNA

Position	Sequence	Large subunit rRNA
a) 2655 -	GUACGAGAGGAC	E. coli (17)
	GUACG$^{A}_{C}$RAGGA$^{G}_{C}$	Consensus
1162 > 1156	GUACGCAAGGAU	T. brucei (13); C. fasciculata
	*	Transition
b)		
2577 -	AGCUGGGUUUA	E. coli (17)
	AG$^{Y}_{G}$U$^{G}_{C}$GGUU$^{UA}_{AC}$	Consensus
1089 > 1082	AGUUGGGUUAA	T. brucei (13); C. fasciculata
c)	2698	
2445 -	GGGAUAACAG---ACCUCGAUGU	E. coli (17)
	RGGAUAACAG---ACCUCG$^{C}_{A}$UGU	Consensus
	1053	
1007 -	AGGUUAACAA---ACCAUGAUUG	T. brucei (13)
	* *	Transitions
	* * **	Transversions
	1044	
998 -	AGGUUAACAA---GCUA$^{AA}_{AA}$AUUG	C. fasciculata
	* * * *	Transitions
	* **	Transversions
	**	Insertions
d) 2430 -	AUAAAAGGU	E. coli (17)
	R$^{UCC}_{AAA}$AAR$^{UU}_{GA}$	Consensus
992 > 989	AUAAAAGGU	T. brucei (13); C. fasciculata
e) 2071 -	ACCUUUAC	E. coli (17)
	CR$_{AC}$CUU$^{U}_{A}$A$^{C}_{A}$	Consensus
895 > 885	UCCUUUAC	T. brucei (13); C. fasciculata
	*	Transitions

TABLE III (continued)

Position	Sequence	Large subunit rRNA
f) 2242 -	GUUUGACUGGGGCG	E. coli (17)
	$\begin{matrix}G\\U\end{matrix}UUY\begin{matrix}RR\\UU\end{matrix}$XYGGGGYR	Consensus
932 / 923	GUUUGAUUGGGGCA	T. brucei (13); C. fasciculata
g) 2053 -	GACGGAAAGACCC	E. coli (17)
	RACGRRAAGACCC	Consensus
878 / 868	UAGGGCAAGUCCU	T. brucei (13); C. fasciculata
	*	Transitions
	* * * *	Transversions

Conserved sequences from trypanosome 12S RNA are shown aligned with their homologues which are universal in the 23S-like rRNAs from the three major kingdoms (typified by E. coli). Consensus sequences take into consideration the corresponding regions from mitochondrial large subunit rRNAs from Paramecium (31); Yeast (46); Aspergillus (52); Maize (48); Human (49); Cow (30); Mouse (50); Mosquito (53).
R = purine; Y = pyrimidine; X = purine or pyrimidine; * indicates positions at which a transition, transversion or insertion occurs in the trypanosome 12S RNA with respect to the consensus sequence; o indicates the position where the aminoacyl end of P-site tRNA has been cross-linked to 23 S rRNA of E. coli (43).

antibiotic resistance/sensitivity of protein synthesis and are in the vicinity of the ribosomal P-site, has about the same size in all organisms and organelles, including trypanosome mitochondria (indicated by the shaded area in the insert of Fig.7). Helices, protruding from this central part, are considerably reduced in size or are eliminated in all mitochondrial large subunit rRNAs. In trypanosome 12S RNA, however, the size reduction of these helices is most extreme.

We have not found extensive structural homologies between the 5'-terminal region of large subunit rRNA and the analogous region of trypanosomal 12S RNA. This is due to the extreme sequence divergence between trypanosome 12S RNA with respect to other rRNAs. The functionally important UAGCUGGUU-811 (E. coli) region, which is conserved in all eubacterial, chloroplast, mitochondrial and eukaryotic cytoplasmic ribosomes is not recognizable in trypanosome 12S RNA. It has been proposed that the 3'-CCA of tRNA, which binds to the ribosome, interacts with the UGG in this conserved sequence (43). Absence of this RNA region in trypanosomal mitoribosomes would suggest that such a mode of tRNA

binding does not occur, at least not in this organelle or that tRNAs may not have CCA at their 3'-termini in trypanosome mitochondria. Since precise positioning of the 3'-ends of tRNA molecules on the ribosome is obviously crucial for the peptidyl transferase reaction, it is possible that in trypanosome mitoribosomes this is accomplished in some other way.

CONCLUDING REMARKS

From comparisons with primary and secondary structures of authentic rRNAs we conclude that the trypanosomal 9S and 12S mitochondrial RNAs are true ribosomal RNAs, albeit strange ones with respect to size, nucleotide sequence, base composition and gene organization. This is attributable in part to a different evolutionary pressure imposed on trypanosomal mtDNA, compared with mtDNA of other organisms. In addition, mitochondrial sequence evolution can be much faster than the evolution of nuclear sequences (33). Nevertheless, a few universal sequences and secondary structure elements have been highly conserved in trypanosome mitochondrial rRNAs. These ancient sequences and structures must therefore be considered as fundamental to the translation process. Universal sequences that are absent in trypanosome 9S and 12S RNAs may represent elements that are either not necessary in trypanosome mitochondrial protein synthesis or for which the trypanosome mitoribosome has developed its own equivalents. The above mentioned unique features of trypanosome mitochondrial rRNAs, the fact that chloramphenicol- or erythromycin-sensitive and cycloheximide-resistant protein synthesis have not been demonstrated in trypanosomes (29), and all attempts to purify ribosomes from trypanosome mitochondria by standard procedures have failed (4,15), lead us to expect that trypanosome mitoribosomes will prove to be most exceptional.

ACKNOWLEDGEMENTS

We appreciate the help of Dr.J.Bol and Mr.F.Brederode from the State University Leiden, The Netherlands and Ms.B.Klaver in 5'-end determination of RNA. We thank Drs. L.A.Grivell and H.F.Tabak for critically reading of the manuscript, Drs.B.F.De Vries for assistance in the preparation of figures 5 and 7 and Ms.T.Scholts for expert typing.
P.S. acknowledges EMBO for a short term fellowship. H.N. thanks K.Katze for discussions, and the U.S.National Institutes of Health (grant no. GM 17129) for support.

*To whom correspondence should be addressed

[1]Present address: Instituto di Fisiologie, Université degli Studi, P.Aldo Moro 5, 0100 Roma, Italy
[2]Present address: The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands

REFERENCES

1. Simpson, L. (1972) Intern.Rev.Cytol. 32, 139-207.
2. Borst, P. and Hoeijmakers, J.H.J. (1979) Plasmid 2, 20-40.
3. Englund, P.T. (1981) in: Biochemistry and Physiology of Protozoa (Levandovsky, M. and Hutner, S.H., Eds), 2nd edn. Vol.4, Academic Press, New York, pp.334-383.
4. Hoeijmakers, J.H.J., Snijders, A., Janssen, J.W.G. and Borst, P. (1981) Plasmid 5, 329-350.
5. Hoeijmakers, J.H.J. and Borst, P. (1978) Biochim.Biophys.Acta 521, 407-411.
6. Simpson, L. and Simpson, A.M. (1978) Cell 14, 169-178.
7. Simpson, A.M. and Simpson, L. (1980) Mol.Biochem.Parasitol. 2, 93-108.
8. Stuart, K.D. and Gelvin, S.B. (1982) Mol.Cell.Biol. 2, 845-852.
9. Simpson, A.M., Simpson, L. and Livingston, L. (1982) Mol.Biochem. Parasitol. 6, 237-252.
10. Borst, P., Hoeijmakers, J.H.J., Hajduk, S.O. (1981) Parasitology 82, 81-93.
11a. Benne, R., De Vries, B.F., Van den Burg, J. and Klaver, B. (1983) Nucl.Acids Res. 11, 6925-6941.
11b. Johnson, B.J.B., Hill, G.C. and Donelson, J.E. (1984) Mol.Biochem. Parasitol. 13, 135-146.
12. Hensgens, L.A.M., Brakenhoff, J., De Vries, B.F., Sloof, P., Tromp, M.C., Van Boom, J.H. and Benne, R. (1984) Nucl.Acids Res. 12, 7327-7344.
13. Eperon, I.C., Janssen, J.W.G., Hoeijmakers, J.H.J. and Borst, P. (1983) Nucl.Acids Res. 11, 105-125.
14. Hoeijmakers, J.H.J., Schoutsen, B., Borst, P. (1982) Plasmid 7, 199-209.
15. Stuart, K. (1983) Mol.Biochem.Parasitol. 9, 93-104.
16. Woese, C.R., Gutell, R., Gupta, R. and Noller, H.F. (1983) Microbiol.Rev. 47, 621-669.
17. Noller, H.F., Kop, J., Wheaton, V., Brosius, J., Gutell, R., Kopylov, A.M., Dohme, F., Herr, W., Stahl, D.A., Gupta, R. and Woese, C.R. (1981) Nucl.Acids Res. 9, 6167-6189.
18. Noller, H.F. (1984) Ann.Rev.Biochem. 53, 119-162.
19. Huysmans, E., Dams, E., Vandenberghe, A. and De Wachter, R. (1983) Nucl.Acids Res. 11, 2871-2880.
20. Auffray, C. and Rougeon, F. (1980) Eur.J.Biochem. 107, 303-314.
21. Birnboim, H.C. and Doly, J. (1979) Nucl.Acids Res. 7, 1513-1523.
22. Smith, H.O. (1980) In: Methods in Enzymology 65 (Grossman, L. and Moldave, K., Eds) Academic Press, New York, pp. 371-380.
23. Poncz, M., Solowiejczyk, D., Ballantine, M., Schwartz, E. and Surrey, S. (1982) Proc. Natl.Acad.Sci.U.S. 79, 4298-4302.
24. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc.Natl.Acad. Sci.U.S. 74, 5463-5467.
25. Van der Marel, G.A., Marugg, J.E., De Vroom, E., Wille, G., Tromp, M., Van Boeckel, C.A.A. and Van Boom, J.H. (1982) Recl.Trav.Chim. Pays-Bas 101, 234-241.
26. Donis-Keller, H., Maxam, A.M., Gilbert, W. (1977) Nucl.Acids Res. 4, 2527-2538.
27. Silberklang, M., Prochiantz, A., Haenni, A.L. and Rajbhandary, U.L. (1977) Eur.J.Biochem. 72, 465-478.
28. Van der Ploeg, L.H.T., Liu, A.Y.C., Michels, P.A.M., De Lange, T., Borst, P., Majumder, H.K., Weber, H., Veeneman, G.H. and Van Boom, J.H. (1982) Nucl. Acids Res. 10, 3591-3604.
29. Benne, R., Agostinelli, M., De Vries, B.F., Van den Burg, J., Klaver, B. and Borst, P. (1983) In: Mitochondria: Nucleo-Mitochondrial Interactions (Schweyen, R.J., Wolf, K. and Kaudewitz, F., Eds) De Gruyter, Berlin, pp.285-302.
30. 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.
31. Seilhamer, J.J., Olsen, G.J. and Cummings, D.J. (1984) J.Biol.Chem. 259, 5167-5172.
32. Seilhamer, J.J., Gutell, R.R., Cummings, D.J. (1984) J.Biol.Chem. 259, 5173-5181.
33. Brown, W.M., Prager, E.M., Wang, A. and Wilson, A.C. (1982) Mol. Evol. 18, 225-239.
34. Maly, P. and Brimacombe, R. (1983) Nucl.Acids Res. 11, 7263-7286.
35. Prince, J.B., Taylor, B.H., Thurlow, D.L., Ofengand, J. and Zimmermann, R.A. (1982) Proc.Natl.Acad.Sci.U.S. 79, 5450-5454.

36. Douthwaite, S., Christensen, A., Garrett, R.A. (1983) J.Mol.Biol. 169, 249-279.
37. Dubin, D.T. and HsuChen, C.C. (1983) Plasmid 9, 307-320.
38. Chan, Y.L., Gutell, R., Noller, H.F. and Wool, I.G. (1984) J.Biol. Chem. 259, 224-230.
39. Van Knippenberg, P.H., Van Kimmenade, J.M.A. and Heus, H.A. (1984) Nucl.Acids Res. 12, 2595-2604.
40. Chan, Y.L., Endo, Y. and Wool, I.G. (1983) J.Biol.Chem. 258, 12768-12770.
41. Schindler, D.G. and Davies, J.E. (1977) Nucl.Acids Res. 4, 1097-1110.
42. Garrett, R.A., Christensen, A. and Douthwaite, S. (1984) J.Mol.Biol. 179, 689-712.
43. Barta, A., Steiner, G., Brosius, J., Noller, H.F. and Kuechler, E. (1984) Proc.Natl. Acad.Sci.U.S. 81, 3607-3611.
44. Monro, R.E., Staehelin, T., Celma, M.L. and Vazquez, D. (1969) Cold Spring Harbor Symp.Quant.Biol. 34, 357.
45. Sor, F. and Fukuhara, H. (1982) Nucl.Acids Res. 10, 6571-6577.
46. Li, M., Tzagoloff, A., Underbrink-Lyon, K. and Martin, N.C. (1982) J.Biol.Chem. 257, 5921-5928.
47. Köchel, H.G. and Küntzel, H. (1981) Nucl.Acids Res. 9, 5689-5696.
48. Chao, S., Sederoff, R. and Levings, C.S. (1984) Nucl.Acids Res. 12, 6629-6644.
49. Eperon, I.C., Anderson, S., Nierlich, D.P. (1980) Nature 286, 460-467.
50. Van Etten, R.A., Wahlberg, M.W. and Clayton, D.A. (1980) Cell 22, 157-170.
51. Kobayashi, M., Seki, T., Katsuyuki, Y. and Koike, K. (1981) Gene 16, 297-307.
52. Köchel, H.G. and Küntzel, H. (1982) Nucl.Acids Res. 10, 4795-4801.
53. HsuChen, C.C., Kotin, R.M., Dubin, D.T. (1984) Nucl.Acids Res. 12, 7771-7785.