Volume 13 Number 7 1985 Nucleic Acids Research

Primary sequence and partial secondary structure of the 12S kinetoplast (mitochondrial) ribosomal RNA from *Leishmania tarentolae:* conservation of peptidyl-transferase structural elements

Vidal F.de la Cruz, Agda M.Simpson, James A.Lake and Larry Simpson*

Biology Department and Molecular Biology Institute, University of California, Los Angeles, CA 90024, USA

Received 28 December 1984; Revised and Accepted 11 March 1985

ABSTRACT

The sequence of the 1173 nt 12S kinetoplast ribosomal RNA from Leishmania tarentolae was determined from the maxicircle DNA sequence, and the 5' and 3' ends localized by primer runoff and S1 nuclease protection experiments. The gene was shown to be free of introns by S1 nuclease analysis. A partial secondary structure model of the 12S RNA molecule is presented which is equivalent in certain respects to the corresponding portions of the Escherichia coli 23S ribosomal RNA model. Domain II of the E. coli model is completely missing in the kinetoplast model with the exception of several phylogenetically conserved stems and one loop. There is a striking conservation of the functionally important peptidyl-transferase region except for the deletion of a few stems and loops. The 12S RNA is the smallest large subunit ribosomal RNA described to date.

INTRODUCTION

The 9S and 12S RNA species are the major stable non-polyadenylated RNAs transcribed from the maxicircle DNA in the single mitochondrion of the kinetoplastid protozoa (for reviews see 1-5). These RNAs have been proposed to be unusually small mitochondrial ribosomal RNAs (rRNAs) (2, 6-8). However, mitochondrial ribosomes have not been isolated from kinetoplastid protozoa, and the unusual properties of small size and order of transcription (large to small rRNA instead of small to large rRNA) have raised questions as to the ribosomal nature of these RNAs. We have recently presented a complete secondary structure model for the kinetoplast 9S RNA (9), based on the Escherichia coli 16S rRNA model (10, 11), which clearly establishes the 9S RNA as the small subunit mitochondrial rRNA of the kinetoplastid protozoa.

In this paper we present the sequence of the 12S mitochondrial RNA from Leishmania tarentolae and a partial secondary structure model that conforms well to portions of the E. coli 23S rRNA

model (12). The secondary structure model includes the E. coli equivalents of domains II and V, and portions of domains IV and VI. The model is supported by several compensatory base pair changes in the Trypanosoma brucei 12S RNA sequence which preserves the base pairing of the proposed helices. A complete secondary structure model for this unusually small large subunit rRNA must await additional cross-species and cross-strain comparisons, as the regions not presented appear to be extremely diverged from the comparable regions of the E. coli model.

MATERIALS AND METHODS

Cloning

Cloning of the 120 DNA fragment, which contains the 12S rRNA gene, and DNA isolation procedures have been described elsewhere (13-15). Subcloning of the insert into M13 mp7, mp8, or mp9 was carried out using standard procedures (16-20). All restriction and DNA modification enzymes were purchased from Bethesda Research Laboratories (BRL), and New England Biolabs. Enzyme reaction conditions were as recommended by the manufacturer.

DNA Sequencing

The entire 120 DNA insert was sequenced (the region encompassing nt 1582-2415 was presented in ref. 9, and the region encompassing nt 2213-6561 was presented in ref. 21) using the dideoxy chain termination method (22-24) with modifications of the gel technology as described by Garoff and Ansorge (25) and de la Cruz et al (21). Universal sequencing primers were used for all clones (New England Biolabs 15-mer). Klenow fragment of DNA polymerase was obtained from Boehringer Mannheim or BRL. Sequence data was compiled using the DB data base management system of Staden (26, 27). The sequence presented here was determined in both orientations and across all restriction enzyme sites, with a minimum of three individual clones for each nt position.

RNA isolation

T. brucei strain 366D cells were grown as the procyclic forms as described (28). Total cell RNA from these cells was isolated according to Chirgwin et al (29) and treated as described (30) with RNase-free DNase I (31). L. tarentolae kinetoplast RNA (kRNA) was isolated as described (30).

Transcript Mapping

The 5' terminal ends of the L. tarentolae and T. brucei 12S RNAs were mapped by the primer-runoff technique (32) and by direct dideoxy sequencing using the RNA as template (33). The 12S RNA primer (17-mer labeled "S4 primer" in Fig. 1) was synthesized according to Ita et al (34). Reverse transcriptase was purchased from BRL. Annealing conditions were: 65° C for 15 min., followed by 37° C for 15 min., and finally room temperature for 20 min. Runoff (primer extension) reactions contained 1-3 ug RNA (total cell RNA in the case of T. brucei; total kinetoplast RNA in the case of L. tarentolae) and 0.5 pmol of S4 primer in a volume of 50 ul.

The 3' end of the L. tarentolae 12S RNA was mapped by S1 nuclease protection analysis (35, 36), using in vitro labeled antisense RNA generated from clones in pSP62-PL (Riboprobe™, Promega Biotec, Inc.) (9, 37, 38).

S1-nuclease analysis for introns

The presence of intervening sequences in the 12S RNA gene was assayed by S1 nuclease protection experiments (35, 36). The 120 insert fragment which contains almost the complete 12S RNA gene (the EcoRI/Sau3A end fragment of the 120 insert; nt 1-1585 in Fig. 1) was gel isolated from a 120/Sau3A digestion. The DNA fragment (0.5 ug) was heat denatured (100° C, 8 min. in water), followed by quick cooling on ice. Total kinetoplast RNA (100 ug) was added and the mixture adjusted to 40 ul in 40 mM NaCl, 27 mM Tris-HCl, pH 8.3, 5 mM $MgCl_2$, and 4 mM DTT. Annealing of the RNA and DNA was carried out as described above. A 20 ul aliquot was diluted with 430 ul S1 nuclease buffer, 10 units of S1 nuclease (BRL) was added and incubated at 37° C for 1 hr. The sample was ethanol precipitated and electrophoresed in a 1.2% non-denaturing agarose gel, blotted onto Schleicher and Schuell BA83 nitrocellulose paper (39) and probed with nick-translated (40) 120 insert.

RESULTS

Sequence of the 12S RNA Gene

The DNA sequence of the first 1665 nt of the 6.56 kb pLtl20 insert from the maxicircle of L. tarentolae is shown in Fig. 1.

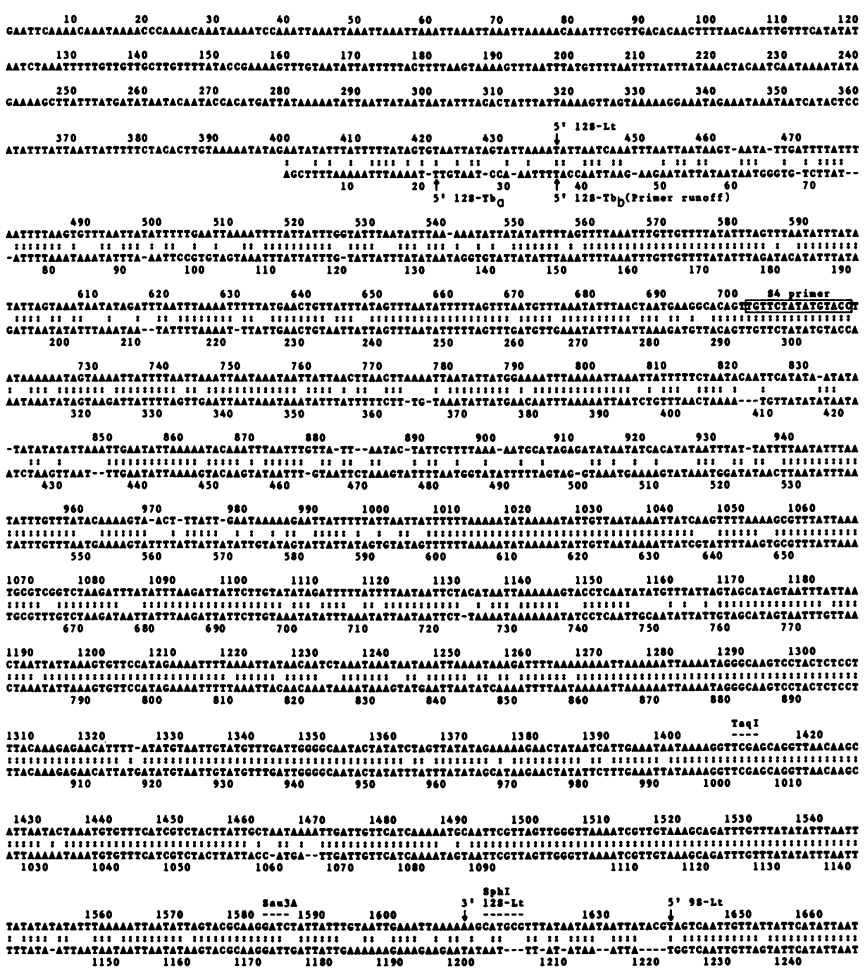

Fig. 1: Nucleotide sequence of the Leishmania tarentolae 12S RNA gene and flanking regions (sense strand; nt 1-1665 of the pLT120 insert).
The first 1665 nt of the L. tarentolae 120 insert sequence in shown. The T. brucei 12S RNA gene sequence (41) is shown below the L. tarentolae sequence in an alignment generated by the Los Alamos SEQA program and modified to accommodate the secondary structure model shown in Fig. 4. Relevant restriction enzyme sites are indicated (see text). The ends of the L. tarentolae 12S and 9S RNAs were determined as described in the text. The S4 primer box is the sequence selected for the oligonucleotide primer complementary to the 12S RNA for use in the 5' end runoff and sequencing experiments. Subscripts: a) 5' end localization of the T. brucei 12S RNA as published by Eperon et al (41). b) 5' end of the T. brucei 12S RNA as determined in this paper (see text).

The published T. brucei 12S RNA gene sequence (41) is shown beneath the L. tarentolae sequence in an alignment determined by computer analysis (Los Alamos SEQA program) and modified slightly to accommodate the secondary structure model shown below.

Localization of the 5' and 3' ends of the 12S RNA

The 5' end of the L. tarentolae 12S RNA was determined by primer runoff using an oligonucleotide primer (labeled "S4 primer" in Fig. 1). The single runoff product was sized on a sequencing gel next to a DNA sequencing ladder of a clone of the 12S gene that was also primed with the S4 primer (Fig. 2A). The 5' end of the RNA was mapped to the location indicated in Fig. 1. Localization of the 5' end was confirmed by direct dideoxy chain termination sequencing using the S4 primer on the RNA template (Fig. 2B). The 5' end of the T. brucei 12S RNA was localized to the position shown in Fig. 1 by primer runoff analysis using the same primer (results not shown; available upon request), the sequence of which is conserved between these two species (41). Our localization for the 5' end of the T. brucei 12S RNA is 15 nt 3' of the end identified by Eperon et al (41).

The 3' end of the L. tarentolae 12S RNA was previously determined by S1 nuclease protection analysis using ^{32}P-labeled antisense RNA transcribed by SP6 RNA polymerase from the pSP62-BC3 hybrid plasmid (9). The 12S RNA protected a 201 nt fragment (see Fig. 3 of reference 9) which represents the distance from the TaqI site (Fig. 1, nt 1409) to the 3' end of the RNA (3' arrow in Fig. 1). A minor 222 nt protected fragment was probably due to 3' end run-on transcription of the 12S RNA gene (in vivo) on the maxicircle DNA. This minor fragment vanished, together with the 201 nt fragment, with the use of template plasmid DNA truncated at the SphI site located at nt 1196 (results not shown, see Fig. 3 of reference 9). We have not verified whether the 3' end of the T. brucei 12S RNA is coincident with the L. tarentolae localization, but it is highly unlikely that the published localization (41) is correct since it overlaps the 5' end of the 9S RNA gene by almost 60 nt (9).

Lack of introns in the 12S RNA

The above S1-protection experiment also established the absence of introns at the 3' end of the 12S RNA gene. The absence of

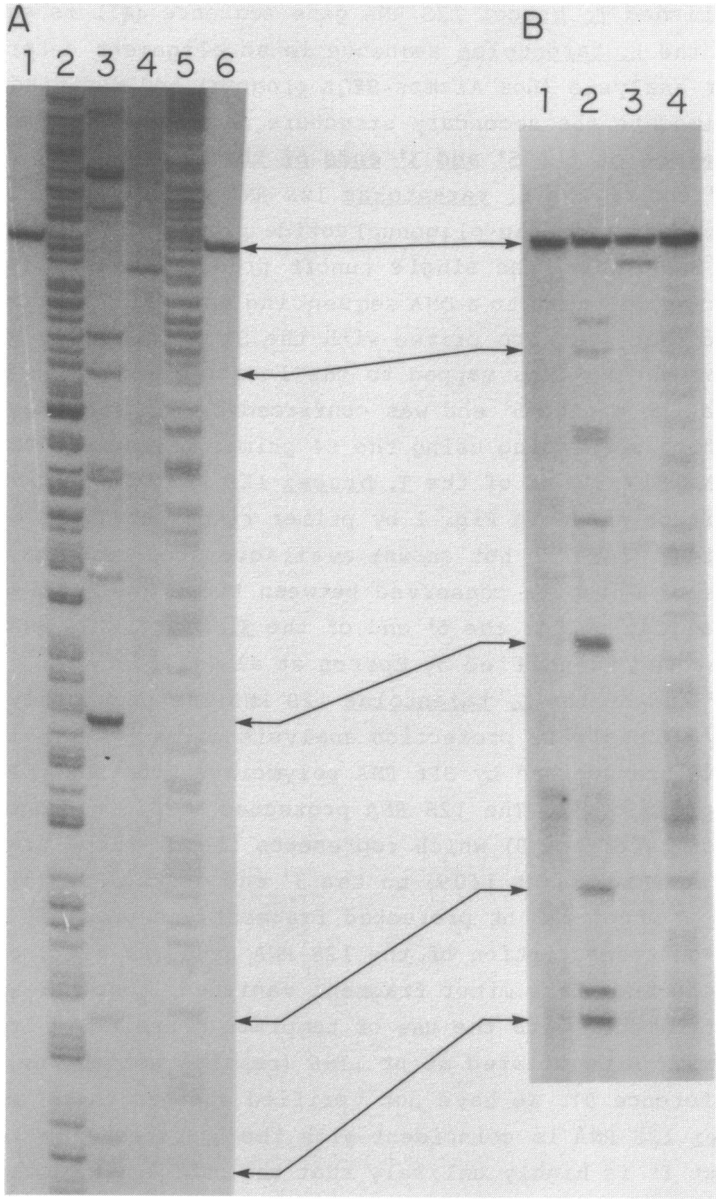

Fig. 2: Determination of the 5' end of the L. tarentolae 12S RNA. A. The 12S RNA 5' end runoff product (lanes 1, 6) co-run with a DNA sequencing reaction (lanes 2, 3, 4, 5 are A, C, G, T reactions, respectively) of a clone containing the sense strand of the 12S RNA gene. The same primer was used on the RNA and the DNA templates so that the 5' terminal nt of the RNA could be read

from the sequencing reaction at the position that comigrates with the 5' runoff product.
B. Dideoxy sequencing reaction using the 12S RNA as template. Although the A and T lanes are not optimal in this experiment, the same sequence ladder pattern was obtained in the C and G lanes as with the DNA clone shown in panel A. Lanes 1, 2, 3, 4 are A, C, G, T reactions, respectively.

introns in the remainder of the gene was shown by the S1 nuclease protection data shown in Fig. 3. The EcoRI/Sau3A fragment 2 of the 120 fragment was hybridized to total kinetoplast RNA and the S1-protected fragment (1139 ± 6 nt) was shown by gel electrophoresis (Fig. 3A) and probe hybridization (Fig. 3B) to correspond in size to full length 12S RNA (expected size is 1144 nt).

The L. tarentolae 12S RNA is extremely A+T-rich (85%) and, in agreement with its putative non-protein coding function, does not exhibit the pronounced strand asymmetry characteristic of maxicircle protein coding sequences in which the T/A ratio is approximately 2.0 (21). The strand asymmetry value (calculated as (|A-T|+|G-C|)/(A+G+C+T)) of the 12S RNA is 0.042 and the T/A ratio is 1.03. Termination codons are frequently found in all six reading frames, also implying that this sequence is not protein coding.

Proposed secondary structure of 12S RNA

The proposed partial secondary structure for the L. tarentolae 12S RNA is shown in Fig. 4 superimposed on the corresponding regions of the E. coli 23S rRNA model (12). We are only presenting the secondary structure for the equivalents of the E. coli 23S rRNA domains II and V, and portions of IV and VI. These regions include: 12S RNA nt 207-299, 508-540, 564-583, and 799-1153 (Fig. 4) (nt 1 of the RNA is nt 438 in Fig.1).

Domain I of 12S RNA The kinetoplast equivalent of domain I has little primary or secondary structural similarity with domain I of the E. coli 23S rRNA. A lack of sequence similarity in this domain has also been observed for the mammalian mitochondrial sequences (42). The domain of the kinetoplast sequence is reduced in size when compared to the E. coli 23S rRNA, as are the mammalian and dipteran mitochondrial sequences (42-44). This

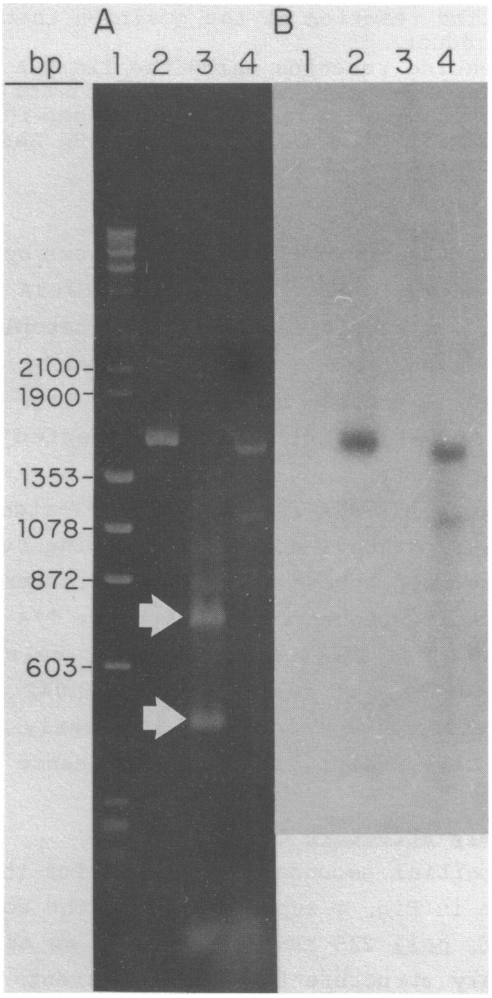

Fig. 3: S1 nuclease protection experiment with L. tarentolae 12S RNA.
A. Ethidium bromide stained non-denaturing 1.2% gel of S1 nuclease protection experiment. Lane: 1, Lambda/HindIII and ØX/HaeIII DNA markers; 2, Gel isolated 120/Sau3A fragment 2 only; 3, kRNA only (This lane was included to show the integrity of the 9S and 12S RNAs, marked with arrows, and the low level of contamination with the cytoplasmic rRNA components; the mobilities of the 9S and 12S RNAs do not correspond to the DNA size markers because this is a non-denaturing gel.); 4, Protected fragments from S1 nuclease digestion of 120/Sau3A fragment 2 with kRNA.
B. Hybridization pattern (probed with the 120 insert) of the gel shown in panel A. Only full size double stranded DNA fragment (1576 nt) and a 1139 (±6) nt nuclease S1-protected fragment were detected.

domain in the 12S RNA is only 31% of the size of the E. coli domain I.

Due to the lack of significant primary sequence similarity in domain I between E. coli and the 12S RNA sequences, a satisfactory secondary structure could not yet be derived for this region. The apparent rapid evolution of this domain, which is apparent upon comparative analysis of many rRNAs, is also evidenced by the increased level of variation between the two kinetoplast sequences: this domain is 71 nt long in L. tarentolae and 67 nt long in T. brucei, with a sequence similarity of only 67%

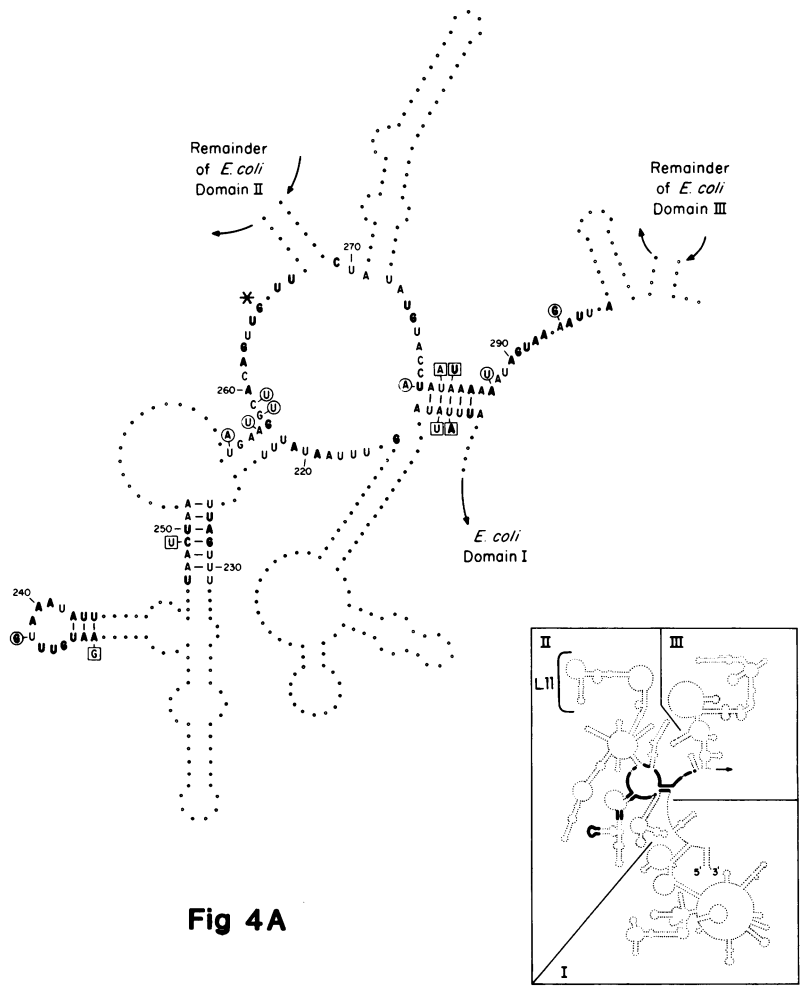

Fig 4A

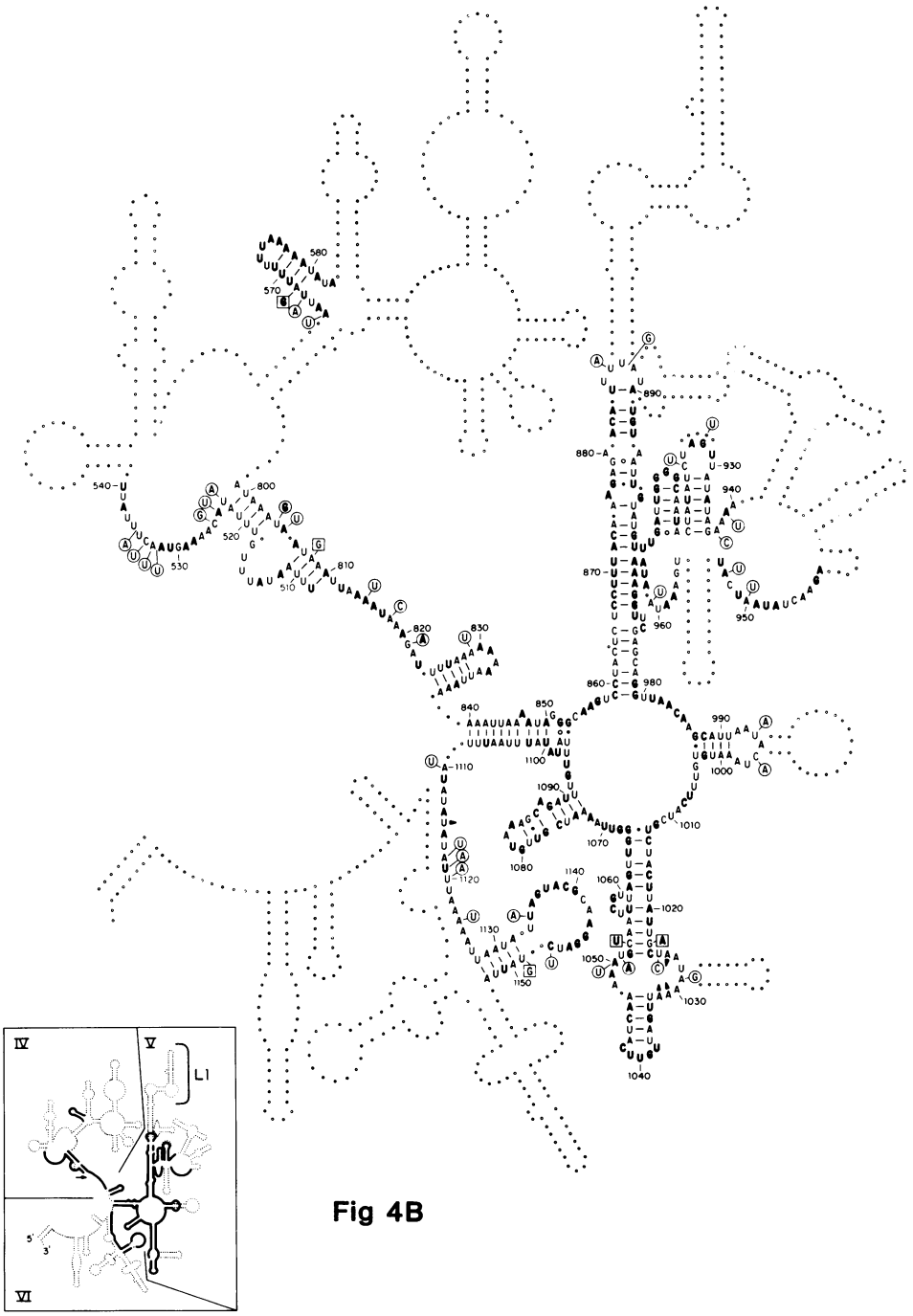

Fig 4B

between the two as compared to a similarity of 81% for the entire 12S RNA.

Domain II of 12S RNA This domain of the kinetoplast 12S RNA (Fig. 4A) is also drastically reduced, representing only 17% of the size of domain II of the E. coli 23S rRNA.

Helix 208-212/281-285 is a long range interaction that bounds the domain. Two compensatory base pair changes between L. tarentolae and T. brucei 12S RNA support this helix (nt 210/283 and 211/282).

Helix 232-234/243-245 and its apex loop show a 12 out of 14 nucleotide match with the corresponding E. coli 23S RNA sequence (E. coli 23S rRNA nt 742-755). This region is strictly conserved in sequence in other large subunit rRNAs, including the mammalian and dipteran mitochondrial rRNAs (42-44).

Helix 225-230/247-252 has been superimposed onto the 687-697/765-775 helix in the E. coli 23S rRNA model in order to connect L. tarentolae helix 232-234/243-245 with the remainder of domain II.

Domain III of 12S RNA Noller et al. (45) and Maly and Brimacombe (10) have noted that domain III is somewhat idiosyncratic. A

Fig. 4: Proposed secondary structure for portions of the L. tarentolae 12S RNA, superimposed on the E. coli 23S rRNA secondary structure model (open dots) of Noller (12).
A. The 5' portion of the 12S RNA. A secondary structure is shown for domain II only. The putative UG- tRNA-binding locus is indicated by an asterisk. The ribosomal protein L11 binding region in E. coli is indicated in the cartoon. The secondary structures for 12S RNA nt 1-206 and 330-507 are not shown pending further cross-species comparisons.
B. The 3' portion of the 12S RNA. The secondary structures for 12S RNA nt 541-563 and 584-798 are not shown pending further analysis. The ribosomal protein L1 binding region in E. coli is indicated in the cartoon.
Nucleotides conserved between the kinetoplast 12S RNA and the E. coli 23S RNA are shown in bold face. The numbers refer to the L. tarentolae 12S RNA sequence, the 5' nt being position one (i.e. position 438 in Fig. 1). Mismatches in the T. brucei 12S RNA sequence (41) are indicated adjacent to the L. tarentolae sequence. Nucleotides within circles are non-compensatory base changes (with regard to base-pairing) in the T. brucei sequence, and nucleotides within boxes are compensatory changes. Deletions in the T. brucei sequence are indicated by solid triangles and insertions are indicated by lines pointing between the appropriate L. tarentolae nucleotides. The boundaries of the six domains of the large subunit rRNA of E. coli (dots) and the 12S RNA (bold lines) are indicated in the cartoons.

generalized structure has not yet been constructed for the prokaryotic, eukaryote cytoplasmic and organellar rRNA models. In agreement with this observation we find that in the case of the kinetoplast 12S RNA there is extensive heterogeneity between the L. tarentolae and T. brucei sequences, similar to that noted for domain I. We do not present a secondary structure for this domain.

Domain IV of 12S RNA Two helices of this domain in the kinetoplast 12S RNA model are shown in Fig. 4B. Other regions of this domain are not shown pending further cross-species sequence comparisons.

Helix 508-522/800-810 is a long range interaction that bounds the domain. The single-stranded region following this helix (12S RNA nt 524-540) tends to be conserved in sequence in other rRNAs, but there is only limited sequence similarity between the kinetoplast 12S RNAs and the E. coli 23S rRNA. The single-stranded region 812-823 is also well conserved.

Helix 568-571/577-580 and the loop bounded by it are maintained in sequence in all rRNAs. Since compensatory base pair changes have not been observed by comparative analysis of other rRNAs, this helix is not considered to be well supported in any model (10). In the 12S RNAs, however, there are some sequence differences that support the existence of at least a portion of this helix.

Domain V of 12S RNA Most of this domain (Fig. 4B) is extremely well conserved in primary sequence and secondary structure in all rRNAs. The central loop, which is thought to play a role in the peptidyl-transferase reaction, is similar in all rRNAs.

Helix 824-828/835-839 is maintained in all large subunit rRNAs. However, the apex loop differs in the kinetoplast 12S RNA by a single transition (AAAA instead of AAGA).

Helix 840-853/1096-1109 is present in all rRNAs. Although there is negligible primary sequence conservation, there is a striking maintenance of the secondary structure and the bulges. The G-A base pair at nt 852/1097, which is present as a G-C base-pair in the E. coli 23S RNA, provides an example of noncanonical base pairing (44).

Helix 859-871/965-979 is extremely constant in structure in all

rRNAs. The ends of the helix tend to be conserved in sequence, while the center of the helix is preserved in secondary structure only. Since this helix is phylogenetically conserved in structure, we would like to point out the possibility of an A-C base pair in the 12S RNA model at nt 862/976, and a U-U base pair at nt 866/970.

Helix 872-884/890-903 represents a region that is much shorter than the corresponding *E. coli* helix, but is very similar to the equivalent regions as proposed in the models for the mammalian mitochondrial (42, 44) and *Aspergillus nidulans* mitochondrial (46) rRNAs. All of these rRNAs appear to be missing the ribosomal protein L1 binding region of *E. coli* (Fig. 4B). There are two possible A-G base pairs in this helix at nt 876/898 and nt 879/895.

Helix 906-910/914-918 represents a phylogenetically conserved helix. The apex loop of this helix is invariant in sequence in all rRNAs. There is one A-G base pair in the 12S RNA model at nt 906/918.

Helix 919-924/931-936 is maintained in all large subunit rRNAs in secondary structure only. This region of the 12S RNA deviates from the equivalent *E. coli* region even more than do the mammalian and dipteran mitochondrial sequences.

The single-stranded regions 854-858, 980-987 and 1004-1012 in the 12S RNA are highly conserved phylogenetically and form important loci for ribosome function (see discussion below).

Helix 987-991/999-1003 is shortened as compared to the *E. coli* 23S rRNA model. The apex loop is strictly maintained in sequence phylogenetically, except in the 12S RNA. This region has been presented in an alternative secondary structure in the *T. brucei* 12S RNA (41), based on primary sequence similarity with the loop in the *E. coli* 23S rRNA. Our model, although containing fewer matches with the *E. coli* sequence, maintains a small number of transversions at the superimposed positions while keeping the central loop of domain V intact.

The complex helix 1013-1037/1043-1067, which is well conserved in all rRNAs, is similarly preserved in the 12S RNA with the exception of a missing side stem. The loop at the apex (nt 1038-1042) is strictly conserved phylogenetically. There are two pos-

sible U-U base pairs (nt 1015/1065, 1019/1061). The T. brucei 12S RNA shows one compensatory base pair change (12S RNA nt 1023/1053, Fig. 4B).

Helix 1073-1079/1084-1091 represents a phylogenetically conserved structure.

Domain VI of 12S RNA Domain VI of the 12S RNA is extremely reduced in size as compared to domain VI of E. coli 23S RNA. This domain is 278 nt long in the E. coli 23S RNA but only 64 nt long in the L. tarentolae 12S RNA. The mammalian mitochondrial rRNAs also show a reduction in the size of this domain when compared with the bacterial sequences, but still contain approximately twice the number of nucleotides for this region than does the 12S RNA. The only phylogenetically conserved region is a loop known to be the site of cleavage by alpha-sarcin (12S RNA nt 1133-1148) (47). The stem in the kinetoplast 12S RNA is conserved only in secondary structure.

DISCUSSION

The selection of specific 12S RNA sequences for construction of a secondary structure model was made on the following basis: (i) The presence of specific rRNA sequences which are highly conserved across widely divergent species lines. (ii) The relative positioning of 12S RNA helices, with respect to each other, in terms of the general E. coli 23S rRNA model. (iii) The construction of equivalent helices in both L. tarentolae and T. brucei 12S RNAs. These criteria are satisfied for the regions of the 12S RNA model presented here. Furthermore, because of the large difference in size between the 12S RNA and the E. coli 23S RNA, we suggest that the regions missing in the 12S RNA are of lower functional importance than those retained, or are regions whose function can be replaced by proteins. Nevertheless, based on the analysis presented here and elsewhere (41), it is clear that the 12S RNA is the mitochondrial large subunit rRNA of the kinetoplastid protozoa.

Overall organization of structure

The 12S RNA model shows a reduction in size of domains I, II, III and VI, when compared to the E. coli 23S rRNA. A similar reduction in the size of domains I, III and VI has been found in

mammalian and dipteran mitochondrial sequences. These domains in the E. coli 23S rRNA have been shown to form a compact structure in a region of the 50S subunit that is RNase T1-resistant and is therefore presumed to have a structural role in the ribosome (42, 48). Since these domains in the 12S RNA have been drastically reduced, this could represent a reordering of an entire region of the large subunit of the ribosome.

Functional sites

Regions of the large subunit RNA secondary structure thought to have functional significance are described below.

The alpha sarcin site Alpha-sarcin inhibits energy utilizing steps of protein synthesis (namely the EF 1-dependent binding of amino-acylated tRNA and the EF 2- and ribosome-dependent hydrolysis of GTP), and cleaves a site in domain VI of the large subunit rRNA (12S RNA nt 1133-1148, Fig. 4B; (49-51)). The single stranded alpha sarcin cleavage site is strictly conserved in primary sequence in all rRNAs, including the kinetoplast 12S RNA. Furthermore, in the 12S RNA the region surrounding this sequence can form a helix such that the sarcin cleavage site is present as a loop, despite lack of sequence similarity to E. coli in the helical region.

The tRNA binding locus The binding of amino-acylated tRNA to the ribosome has been proposed to occur via a base-pairing interaction of the terminal CCA of the tRNA to a specific UGG sequence on the large subunit rRNA (52). If a common mechanism is at work in all ribosomes, then a UGG at an equivalent structural locus should be found in all large subunit rRNAs. The site proposed for this possible base-pairing of tRNA and rRNA in E. coli is 23S RNA nt 808-810 (52). This site is conserved in all rRNAs analyzed to date, but not in the kinetoplast 12S RNA, where a UGU sequence is present (nt 265-267, Fig. 4A). It is possible that there is no binding of tRNA in the kinetoplast ribosome via base pairing at this site, or, alternatively, that the base pairing of two nucleotides (i.e. the nucleotides closest to the bound amino acid residue) is sufficient for tRNA binding.

The puromycin binding locus Puromycin, which mimics amino-acylated tRNA, causes release of nascent peptides from tRNA bound at the P site of the ribosome (51). A derivative of puromycin has

been cross-linked to E. coli 23S rRNA (53, 54), and the location of the cross-linked RNA sequence has been mapped to E. coli 23S rRNA loop 2552-2556 (42). This loop is phylogenetically conserved and is present in the kinetoplast 12S RNA (12S RNA nt 1038-1042, Fig. 4B). The sequence in the 12S RNA model is located in an equivalent structural environment as compared to the same sequence in the E. coli 23S rRNA model. These observations suggest that this loop plays a direct role in translation. Eperon et al (41) proposed an alternative secondary structure for this region in the T. brucei 12S RNA, lacking the long complex helix 1013-1037/1043-1067 shown in Fig. 4B. Our data, however, appear to favor the presence of this helix.

The peptidyl transferase center There is evidence from affinity labeling experiments and from isolation of mutants resistant to peptidyl-transferase related antibiotics, such as chloramphenicol (Cap) and erythromycin, that the central ring structure in domain V of the E. coli 23S rRNA model is involved with the peptidyl-transferase reaction (12). Five positions in this region have been identified, the modification of which can result in Cap-resistance in yeast (55, 56), and mouse and human (57-59) mitochondria. These positions correspond to nt 979, 983, 984, 1010 and 1011 in the 12S RNA (Fig. 4B). Of these, two positions in the 12S RNA (12S RNA nt 1010, 1011) differ from the E. coli wild type sequence. Although mitochondrial protein synthesis in trypanosomes has not been conclusively demonstrated (60-64), this altered sequence may imply the Cap-resistance of the mitochondrial protein synthesis machinery in trypanosomes. In addition, there is a single nucleotide position at an equivalent structural location in several large subunit rRNAs which is implicated in resistance to erythromycin, macrolide, lincosamide and streptogramin type B antibiotics (65-67). This position in the 12S RNA sequence (nt 854) is a C instead of the A found in the E. coli wild type sequence.

Other conserved regions We have identified in the 12S RNA several other regions with similarity to E. coli structures. These regions are of unknown function, but are clearly conserved phylogenetically in primary sequence and secondary structure, and have also been proposed to lie near the peptidyl-transferase region in

native ribosomes.

The first region is located in domain II. In the mammalian mitochondrial rRNAs, the most conserved regions in this domain include the equivalents of 12S RNA helices 208-212/281-285 and 225-230/247-252 (Fig. 4B), which lead from the single stranded RNA regions outside this domain (i.e. RNA linking the domains, e.g. 12S RNA nt 288-299), to the equivalent of the *E. coli* 740-757 stem and loop (12S RNA nt 232-245), and the ribosomal protein L11 binding region (nt 1051-1108 in *E. coli* 23S RNA). The L11 binding region (Fig. 4A) is missing in the 12S RNA. We therefore speculate that the binding of the ribosomal protein L11 equivalent in the kinetoplast ribosome is mediated by protein interaction only. The function of the phylogenetically conserved helix 232-234/243-245 (12S RNA, Fig. 4A) is unknown, but in the *E. coli* ribosome this region of the 23S rRNA has been proposed to interact with an RNA sequence in the peptidyl-transferase region via a tertiary interaction (68). The tRNA locus described above is also located in this region (12S RNA nt 265-267, Fig. 4A). It is of interest to note that these helices and phylogenetically conserved stem and loop in the mammalian mitochondrial rRNAs are the only elements maintained in domain II of the kinetoplast 12S RNA.

The fact that other structures in domain II are eliminated in the 12S RNA suggests that these missing regions do not play a crucial role in translation. In agreement with this is the observation that in the mammalian mitochondrial rRNAs these regions are reduced in size and have little primary sequence similarity to the *E. coli* 23S RNA. Furthermore, in domain II of the *Aspergillus nidulans* mitochondrial rRNA, there are several insertions, instead of deletions (46).

Helix 568-571/577-580 (domain IV, Fig. 4b) and its apex loop represent another phylogenetically conserved structure which in the *E. coli* ribosome has been proposed to have a tertiary interaction with the peptidyl-transferase region (68). This putative tertiary interaction and the strict maintenance of primary sequence suggests an important function for this region in translation.

Finally, we note that helix 906-910/914-918 (domain V; Fig. 4B)

and its apex loop are phylogenetically conserved. No function has yet been suggested for this region.

In summary, comparative analysis is a powerful approach to the understanding structure-function relationships of rRNAs, and study of unusual rRNAs such as the 12S RNA may play an important role in this process. Many of the sequences and structures of the *E. coli* 23S rRNA model that are implicated in the biosynthetic function of the ribosome have been identified in the kinetoplast 12S RNA. In addition, several large portions of RNA are not present in the 12S RNA, which suggests that these regions are not crucial for translation. Although this has been previously noted for other mitochondrial rRNAs, the elimination of the majority of domain II is a novel finding. In view of the small size of the 12S rRNA, we speculate that some of the functions of the rRNA have been taken over by ribosomal proteins in the kinetoplast ribosome. These unusually small large subunit rRNAs may be instructive in indicating regions of RNA that are absolutely crucial for basic translational functions.

ACKNOWLEDGEMENTS

These studies were supported in part by an NIH grant to L.S. (AI09102) and by NIH and NSF grants to J.L. (PCM23016, GM24034). V.C. was a predoctoral fellow on NIH training grant GM7104.

*To whom correspondence should be addressed

REFERENCES

1. Simpson, L. (1972) Inter. Rev. Cytol., 32, 139-207.
2. Borst, P. and Hoeijmakers, J. (1979) Plasmid, 2, 20-40.
3. Englund, P. (1981) in Biochemistry and Physiology of Protozoa, Levandowsky, M. and Hutner, S.H. Eds., Vol. 4, PP. 333-383, Academic Press, New York.
4. Englund, P., Hajduk, S. and Marini, J. (1982) Ann. Rev. Biochem., 51, 695-726.
5. Sederoff, R.R. (1984) Adv. Genet., 22, 1-108.
6. Simpson, A.M. and Simpson, L. (1980) Mol. Biochem. Parasitol., 2, 93-108.
7. Hoeijmakers, J., Schoutsen, B. and Borst, P. (1982) Plasmid, 7, 199-209.
8. Simpson, L., Simpson, A. and Livingston, L. (1982) Mol. Biochem. Parasitol., 6, 237-257.
9. de la Cruz, V.F., Lake, J.A., Simpson, A.M. and Simpson, L. (1985) Proc. Nat. Acad. Sci. USA, In press.

10. Maly, P. and Brimacombe R. (1983) Nucl. Acids Res., 11, 7263-7286.
11. Woese, C., Gutell, R., Gupta, R. and Noller, H. (1983) Microbiol. Rev., 47, 621-669.
12. Noller, H.F. (1984) Ann. Rev. Biochem., 53, 119-162.
13. Masuda, H., Simpson, L., Rosenblatt, H. and Simpson, A.M. (1979) Gene, 6, 51-73.
14. Simpson, L. (1979) Proc. Natl. Acad. Sci. USA, 76, 1585-1588.
15. Simpson, L., Simpson, A., Kidane, G., Livingston, L. and Spithill, T. (1980) Am. J. Trop. Med. Hyg.,Suppl., 29, 1053-1063.
16. Sanger, F., Coulson, A., Barrell, B., Smith, A. and Roe, B. (1980) J. Mol. Biol., 143, 161-178.
17. Messing, J., Crea, R. and Seeburg, P.H. (1981) Nuc. Acids Res., 9, 309-321.
18. Winter, G. and Fields, S. (1980) Nucl. Acids Res., 8, 1965-1974.
19. Anderson, S. (1981) Nucl. Acid Res., 9, 3015-3027.
20. Hong, G.F. (1982) J. Mol. Biol., 158, 539-549.
21. de la Cruz, V.F., Neckelmann, N. and Simpson, L. (1984) J. Biol. Chem., 259, 15136-15147.
22. Sanger, F. Nicklen, S. and Coulsen, A.R. (1977) Proc. Natl. Acad. Sci. USA, 74, 5463-5467.
23. Sanger, F. and Coulson, A.R. (1978) FEBS Lett., 87, 107-110.
24. Hong, G.F. (1981) Biosci. Rep., 1, 243-252.
25. Garoff, H. and Ansorge, W. (1981) Anal. Biochem., 115, 450-457.
26. Staden, R. (1980) Nucl. Acid Res., 8, 3673-3694.
27. Staden, R. (1982) Nucl. Acids Res., 10, 4731-4751.
28. Simpson, L. (1978) J. Parasitol., 64, 360.
29. Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W. J. (1979) Biochem., 18, 5294-5299.
30. Simpson, L. and Simpson, A. (1978) Cell, 14, 169-178.
31. Zimmerman, S. and Sandeen, G. (1966) Anal. Biochem., 14, 269-277.
32. Proudfoot, N., Shander, M., Manley, J., Gefter, M. and Maniatis, T. (1980) Science, 209, 1329-1336.
33. Sures, I., Levy, S. and Kedes, G. (1980) Proc. Nat. Acad. Sci. USA, 77, 1265-1269.
34. Ita, H., Ike, Y., Ikuta, S. and Itakura, K. (1982) Nucl. Acids Res., 9, 2989-2998.
35. Berk, A. and Sharp, P. (1977) Cell, 12, 721-732.
36. Berk, A.J. and Sharp, P.A. (1978) Proc. Nat. Acad. Sci. USA, 75, 1274-1278.
37. Green, M., Maniatis, T. and Melton, D. (1983) Cell, 32, 681-694.
38. Melton, D.A., Krieg, P.A., Rebagliati, M.R., Maniatis, T., Zinn, K. and Green, M.R. (1984) Nucl. Acids Res., 12, 7035-7056.
39. Smith, G. and Summers, M. (1980) Anal. Biochem., 109, 123-129.
40. Rigby, P.W., Dieckmann, M., Rhodes, C. and Berg, P. (1977) J. Mol. Biol., 113, 234-251.
41. Eperon, I., Janssen, J., Hoeijmakers, J. and Borst, P. (1983) Nucl. Acids Res., 11, 105-125.

42. Branlant, C., Krol, A., Machatt, M.A., Pouyet, J., Ebel, J.-P., Edwards, K. and Kossel, H. (1981) Nucl. Acids Res., 9, 4303-4324.
43. HsuChen, C.-C., Kotin, R.M. and Dubin, D.T. (1984) Nucl. Acids Res., 12, 7771-7785.
44. Glotz, C., Zwieb, C. and Brimacombe, R. (1981) Nucl. Acids Res., 9, 3287-3306.
45. Noller, H.F., Kop, J., Wheaton, V., Brosius, J., Gutell, R.R., Kopylov, A.M., Dohme, F., Herr, W., Stahl, D.A., Gupta, R. and Woese, C.R. (1981) Nucl. Acids Res., 9, 6167-6189.
46. Kochel, H.G. and Kuntzel, H. (1982) Nucl. Acids Res., 10, 4795-4801.
47. Endo, Y. and Wool, I.G. (1982) J. Biol. Chem., 257, 9054-9060.
48. Branlant, C., SriWidada, J., Krol, A. and Ebel, J.-P. (1977) Eur. J. Biochem., 74, 155-170.
49. Fernandez-Puentes, C. and Vasquez, D. (1977) FEBS Lett., 78, 143-146.
50. Schindler, D.G. and Davies, J.E. (1977) Nucl. Acids Res., 4, 1097-1110.
51. Hobden, A.N. and Cundliffe, E. (1978) Biochem. J., 170, 57-61.
52. Barta, A., Steiner, G., Brosius, J., Noller, H.F. and Kuechler, E. (1984) Proc. Nat. Acad. Sci. USA, 81, 3607-3611.
53. Greenwall, P., Harris, R.J. and Symons, R.H. (1974) Eur. J. Biochem., 49, 539-554.
54. Eckermann, D.J. and Symons, R.H. (1978) Eur. J. Biochem., 82, 225-234.
55. Brosius, J., Dull, T.J. and Noller, H.F. (1980) Proc. Nat. Acad. Sci. USA, 77, 201-204.
56. Dujon, B. (1980) Cell, 20, 185-197.
57. Blanc, H., Adams, C.W. and Wallace, D.C. (1981) Nucl. Acids Res., 9, 5785-5795.
58. Blanc. H., Wright, C.T., Bibb, M.J., Wallace, D.C. Clayton, D.A. (1981) Proc. Nat. Acad. Sci. USA, 78, 3789-3793.
59. Kearsey, S.E. and Craig, I.W. (1981) Nature, 290, 607-608.
60. Laub, R. and Thirion, J. (1971) Arch. Internat. Physiol. Biochim., 80, 199-200.
61. Laub-Kupersztejn, R. and Thirion, J. (1974) Biochim. Biophys. Acta, 340, 314-322.
62. Hanas, J., Linden, G. and Stuart, K. (1975) J. Cell Biol., 65, 103-111.
63. Kleisen, C. and Borst, P. (1975) Biochim. Biophys. Acta, 407, 473-478
64. Spithill, T., Shimer, S., and Hill, G. (1981) Mol. Biochem. Parasitol., 2, 235-256.
65. Sor, F. and Fukuhara, H. (1982) Nucl. Acids Res., 10, 6571-6578.
66. Sigmund, C., Ettayebi and Morgan, E. (1984) Nucl. Acids Res., 12, 4653-4664.
67. Skinner, R., Cundliffe, E. and Schmidt, F.J. (1983) J. Biol. Chem., 258, 12702-12706.
68. Stiege, W., Glotz, C. and Brimacombe, R. (1983) Nucl. Acids Res., 11, 1687-1706.