438–446 Nucleic Acids Research, 1997, Vol. 25, No. 2

Identification of additional rRNA fragments encoded by the *Plasmodium falciparum* 6 kb element

Jean E. Feagin[1,2,*], Barbara L. Mericle[1], Erica Werner[1] and Mark Morris[1]

[1]Seattle Biomedical Research Institute, 4 Nickerson Street, Seattle, WA 98109-1651, USA and [2]Department of Pathobiology, School of Public Health and Community Medicine, University of Washington, Seattle, WA 98195, USA

Received May 6, 1996; Revised and Accepted November 22, 1996

ABSTRACT

Sequences similar to mitochondrial large and small subunit rRNAs are found as small scattered fragments on a tandemly reiterated 6 kb element in the human malaria parasite *Plasmodium falciparum*. The rDNA sequences previously identified include strongly conserved portions of rRNA, suggesting that fragmented rRNAs derived from them are able to associate into functional ribosomes. However, sequences corresponding to other expected rRNA regions were not found. We here report that 10 of the 13 previously described rDNA regions have abundant small transcripts. An additional 10 transcripts were found from regions not previously known to contain genes. Five of the latter have been identified as rRNA fragments, including those corresponding to the 5′ end and 790 loop sequences of small subunit rRNA and the sarcin/ricin loop of large subunit rRNA. Demonstration that most of the previously described rDNA regions have abundant transcripts and the identification of new transcripts with other portions of conventional rRNAs provide support for the hypothesis that these small transcripts comprise functional rRNAs.

INTRODUCTION

Mitochondrial protein coding genes are present on tandemly reiterated copies of a 6 kb sequence element in *Plasmodium* species (1–3), thus identifying this element as the malaria parasite's mitochondrial genome. Ribosomal RNAs are also invariably encoded by mitochondrial genomes (4) and the 6 kb element has small regions of sequence similar to rRNAs (2,3,5,6). Unexpectedly, these small rDNA regions are not contiguous: they are interspersed with each other, are scrambled in order, and are encoded on both strands of the DNA. However, the nucleotide sequences are highly conserved between *Plasmodium* species, they correspond to highly conserved regions of rRNA, and at least some have relatively abundant small transcripts (5). These characteristics suggest that rRNA fragments corresponding to these regions may assemble into functional ribosomes.

Fragmented rRNAs, while uncommon, have been described from other eukaryotes (7,8), prokaryotes (9,10), and organelle genomes (11–15). In some cases the fragmentation is minor: the *Tetrahymena pyriformis* mitochondrial rRNAs each have two fragments (11,16). More complex fragmentation is seen for the nucleo-cytoplasmic large subunit rRNAs of trypanosomatid protozoa, which consist of seven fragments (7,17,18), and that of *Euglena gracilis*, which has 14 fragments (8,19). Many examples of fragmented rRNAs share the characteristic that the rRNA fragments are encoded in the expected linear order. The fragmented mitochondrial rRNAs of *T.pyriformis* (11,16) and of *Chlamydomonas* species (12,20), on the other hand, are encoded out of expected order. These examples of fragmented rRNAs all differ significantly from those of *Plasmodium* in their aggregate size. Thirteen rDNA regions, from 28 to 178 nucleotides (nt), have been identified from the *Plasmodium falciparum* 6 kb element, based on significant similarity in nucleotide sequence and potential secondary structure to highly conserved regions of other rRNAs (5). The combined size of regions corresponding to large subunit (LSU) rRNA is 822 nt and for the small subunit (SSU) rRNA, the predicted size total is 429 nt (5). This is much smaller than other fragmented rRNAs, which generally approximate the size of their non-fragmented counterparts. As a comparison, the LSU and SSU rRNAs of *E.coli* rRNAs are 2904 and 1542 nt, respectively, and the corresponding fragmented mitochondrial rRNAs of *Chlamydomonas reinhardtii* sum to 2419 and 1200 nt (12).

Comparison of the *P.falciparum* rRNA fragments to the sequence and secondary structures expected for organelle rRNAs has shown that they correspond to a number of highly conserved regions of rRNA (5). However, sequences corresponding to some expected regions were not found in initial analyses. These include the sarcin/ricin loop of the LSU rRNA and portions at the 5′ end and the 790 loop of the SSU rRNA. It seems unlikely that a ribosome would function without at least some of these rRNA sequences. In addition, although a variety of small transcripts have been shown to hybridize to 6 kb element probes (3,6,21), only a few (5,22) have been shown to correspond to specific rDNA regions. We have employed transcript mapping approaches to verify transcription of the previously described *P.falciparum* rDNA regions and to search for potential rRNA fragments from regions of the 6 kb element not known to encode genes. Our analyses have shown that 10 of the 13 originally proposed rDNA regions encode small, abundant transcripts while similarly abundant transcripts could not be detected for three rDNA regions. We have also detected an additional 10 small, abundant transcripts, five of which have sequence similarity to regions of rRNAs from other organisms and/or contain sequences complementary to the other rRNA fragments.

*To whom correspondence should be addressed at: Seattle Biomedical Research Institute, 4 Nickerson Street, Seattle, WA 98109-1651, USA. Tel: +1 206 284 8846; Fax: +1 206 284 0313; Email: feagin@u.washington.edu

METHODS

Parasites

The C10 line of *P.falciparum* was employed for these studies. Parasites were cultivated by the method of Trager and Jensen (23) and prepared for RNA extraction by lysing infected erythrocytes with saponin, followed by washes with phosphate-buffered saline. Isolated parasites were quick-frozen in liquid N_2 and stored at –80°C for later use.

RNA preparation, blotting and hybridization

Total RNA was prepared from mixed populations of parasites, generally 50% or more trophozoites, by lysis in guanidinium thiocyanate and extraction with acidic phenol:chloroform (24), as previously described (25). For RNA blots, total RNA was electrophoresed on 12% acrylamide, 7 M urea gels in TBE (0.1M Tris-borate, 0.9 mM EDTA) and electrophoretically transferred to nylon membrane in TAE (40 mM Tris-acetate, 1 mM EDTA). Probing with radiolabeled *in vitro* transcripts or oligonucleotides was carried out as described previously (25), with modification to the temperatures for oligonucleotide probings as dictated by their characteristics. Probe locations are shown in Table 1.

Table 1. Probes for RNA blots

Gene	Probe type	Probe location[a]	Transcript size (nt)
SSUA	oligonucleotide	1919–1943	140
SSUB	oligonucleotide	431–452	110
SSUD	oligonucleotide	5395–5417	65
SSUE	oligonucleotide	1656–1675 rc	40
SSUF	oligonucleotide	5468–5489	74, 58
LSUA	*in vitro* transcript	4957–5163	175
LSUD	*in vitro* transcript	5525–5967	78
LSUE	*in vitro* transcript	5525–5967	190
LSUF	*in vitro* transcript	1411–1657 rc	125, 110
LSUG	oligonucleotide	332–353	115
RNA1	*in vitro* transcript	544–664	95
RNA2	*in vitro* transcript	1658–1851 rc	75
RNA3	*in vitro* transcript	1852–2159 rc	85
RNA4	*in vitro* transcript	4526–4748 rc	70
RNA5	*in vitro* transcript	4749–4956 rc	92
RNA6	*in vitro* transcript	4790–4956 rc	58
RNA7	oligonucleotide	5226–5252	94
RNA8	*in vitro* transcript	5894–5967	115
RNA9	*in vitro* transcript	2–130 rc	53
RNA10	oligonucleotide	644–665	100

[a]All locations are given according to GenBank entry M76611. rc, reverse complement of the indicated sequence.

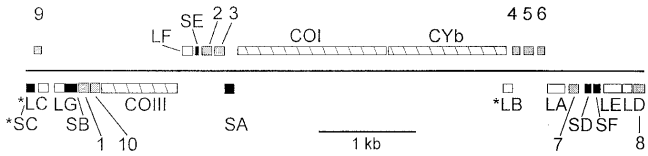

Figure 1. Map of the *P.falciparum* 6 kb element. A schematic map of the 6 kb element is shown, with genes above or below the line depending on direction of transcription (left to right above the line). Because the 6 kb element is tandemly repeated, the endpoints shown reflect random choice rather than actual structure; they correspond to the endpoints of GenBank submission M76611. Protein coding genes are indicated by hatched boxes, previously identified (5) rDNA sequences are shown with white (LSU rRNA) and black (SSU rRNA) boxes and are designated with letters, and the new small RNAs are shown with gray boxes and designated with numbers. The three rDNA regions for which no transcripts were found are marked with asterisks. COI and COIII, cytochrome *c* oxidase subunits I and III; CYb, cytochrome *b*; LA-LG, LSU rRNA fragments; SA-SF, SSU rRNA fragments.

Sequence alignments and potential secondary structures

Sequence data for the *P.falciparum* 6 kb element is found in Genbank entry M76611. The additional rRNAs described here were identified by searching, using the Generunner program (Hastings Software, Inc.), for short highly conserved motifs in the regions known to encode transcripts and then manually assessing the potential for longer alignments. Proposed secondary structures were modeled after consensus structures (26,27) which include rRNAs from three phylogenetic domains (archaea, bacteria, and eukarya) and two organelle types (chloroplasts and mitochondria).

RESULTS

Transcript mapping

The *P.falciparum* 6 kb element is densely populated with genes, having three protein coding genes and 13 predicted rDNA fragments (Fig. 1). To verify that transcripts existed for each rDNA region and to determine whether other potential rRNA fragments were encoded by other portions of the 6 kb element, we embarked on a systematic transcript mapping project. Radiolabeled *in vitro* transcripts were made from short clones, among them covering the entire length of the 6 kb element separately for each strand, and used to probe RNA blots prepared from denaturing acrylamide gels. When necessary, oligonucleotide probes were used to resolve whether multiple transcripts detected by a single *in vitro* transcript probe were derived from the same or adjacent sequences. Abundant small transcripts were found for 10 of the originally proposed rDNA regions (Fig. 2, Table 1) but not for the SSUC, LSUB and LSUC regions, discussed below. Two of the rDNA regions, SSUF and LSUF, have abundant transcripts of two different sizes, as confirmed by hybridization with oligonucleotide probes.

In addition to transcripts which correspond to predicted rRNA fragments, an additional 10 small transcripts were found, ranging from ~50 to 115 nt (Fig. 3, Table 1). The previously identified 6 kb element genes are associated in two clusters, one for each strand. With the exception of RNA9, the new transcripts map within or immediately adjacent to the clusters (Fig. 1). All the transcripts appear relatively abundant. Although we have not closely examined the stoichiometry of the small RNAs, use of

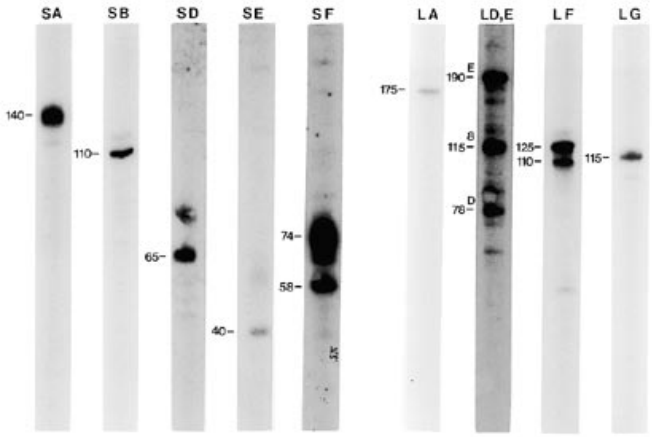

Figure 2. RNA blots of rDNA regions from the *P.falciparum* 6 kb element. Total *P.falciparum* RNA was electrophoresed on denaturing 7 M urea, 12% acrylamide gels, electrophoretically transferred to nylon membrane, and probed with sequences (Table 1) complementary to rDNA regions from the *P.falciparum* 6 kb element (5). Transcript sizes were determined using a ladder of small *in vitro* transcripts as markers. The identities of LSUD (D) and LSUE (E) transcripts, shown on the same panel (LD,E), were later established with oligonucleotide probes. The 115 nt transcript from that panel corresponds to one of the new transcripts, RNA8. Abbreviations are as shown for Figure 1.

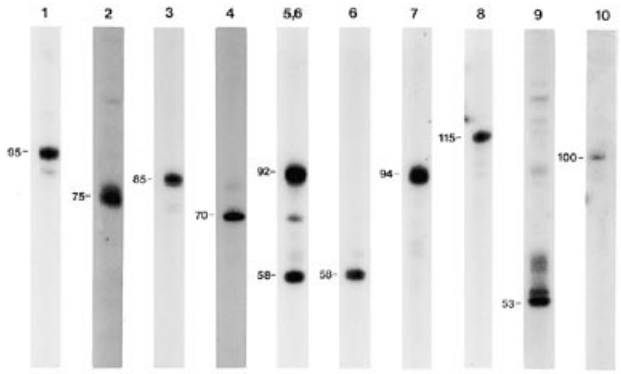

Figure 3. RNA blot analysis of transcripts from 6 kb element sequences lacking identified genes. Blots prepared as for Figure 2 were probed with sequences from regions of the *P.falciparum* 6 kb element which lacked identified genes (Table 1). The numbering of the panels corresponds to the numbered boxes shown in Figure 1. Size markers were as described for Figure 2.

single probes which detect transcripts from more than one gene (Fig. 2, lane LD,E; Fig. 3, lane 5,6) results in similar signal strength for the different RNAs, and suggests that the transcripts are generally similar in abundance.

Because of the highly conserved character of portions of rRNA sequence, it is theoretically possible that some of the small transcripts detected could derive from other rRNAs rather than the 6 kb element sequences used as probes. For *Plasmodium*, this includes both nucleo-cytoplasmic rRNAs and those encoded by a second organelle DNA, the 35 kb DNA (25,28). The region most conserved between the two *P.falciparum* organelle DNAs is the peptidyltransferase domain of LSU rRNA (5). Probes for the corresponding 6 kb element rDNA sequences cross-hybridize to a minor degree with LSU rRNA from the 35 kb DNA but do not cross-hybridize detectably with the nucleo-cytoplasmic LSU rRNA (data not shown). If the most well-conserved sequences do not cross-hybridize strongly, it is unlikely that the abundant transcripts we detect derive from specific breakdown of other rRNAs. In confirmation, probes which show no detectable cross-hybridization to other rRNAs have been defined for 10 of the small transcripts thus far.

Identification of new rRNA fragments

Comparison of rRNA sequences from a wide variety of organisms provides an insight into those regions most strongly conserved. Two types of conservation should be considered. One is conservation of nucleotide sequence and the other is positional conservation, meaning the maintenance of particular structures, such as the size and location of loops and length of helices. Mitochondrial rRNA sequences and potential structures often tend to be unusual. Adding mitochondrial sequences to comparisons of rRNAs from the three phylogenetic domains (archaea, bacteria and eukarya) and chloroplasts produces a noticeable decrease in both sequence and positional conservation (26,27). Consideration of the fragmented *P.falciparum* mitochondrial rRNAs indicates that many correspond to regions with significant sequence and positional conservation (Fig. 4) and have proposed secondary structures consistent with other rRNAs (5). In attempting to identify additional rRNA fragments from the 6 kb element, we first searched for conserved sequence motifs or for sequences complementary to appropriate portions of previously identified rRNA fragments. This was followed by consideration of the potential secondary structure, and its similarity to consensus structures.

New SSU rRNA fragments

Two of the new rRNAs correspond to regions missing from the originally proposed SSU rDNA fragments. One of these, the 70 nt RNA4, was detected with a probe that contains sequence similar to the most conserved portion of the 790 loop of the SSU rRNA. An oligonucleotide complementary to the conserved sequence also hybridizes to a 70 nt transcript (data not shown). The consensus sequence at the terminal loop is GAUNAGAUACC (where N is less highly conserved than the other sites). Ten of the sites in the terminal loop are conserved in >90% of SSU rRNAs in the three domain, two organelle consensus (Fig. 4). Of these, RNA4 matches eight, with its loop sequence being GAUAAGAU-GAC (Fig. 5). The GA at positions 26 and 27 (Fig. 5) replaces the normally conserved AC. The consensus stem below the loop has only a few scattered 90+% conserved nucleotides but a specific structure, with conserved side bulges (Fig. 4). The proposed RNA4 structure conforms very well to the consensus, with similarly placed and sized side bulges (except for the bulged G at position 35, which is missing in the consensus structure) and similar numbers of paired nucleotides in helices.

The search for the 6 kb element sequences corresponding to the 5′ end of SSU rRNA relied on complementarity to the SSUA and SSUB RNAs. Portions of these are expected, based on SSU rRNA secondary structures, to form helices with the 5′ end sequences. A computer search for 6 kb element sequences complementary to the relevant portion of the previously proposed SSUA secondary structure (5) yielded several possible, though imperfect, matches. However, none of these corresponded to regions with small transcripts. The choice of SSUA sequences to include in the search is dictated by the proposed secondary

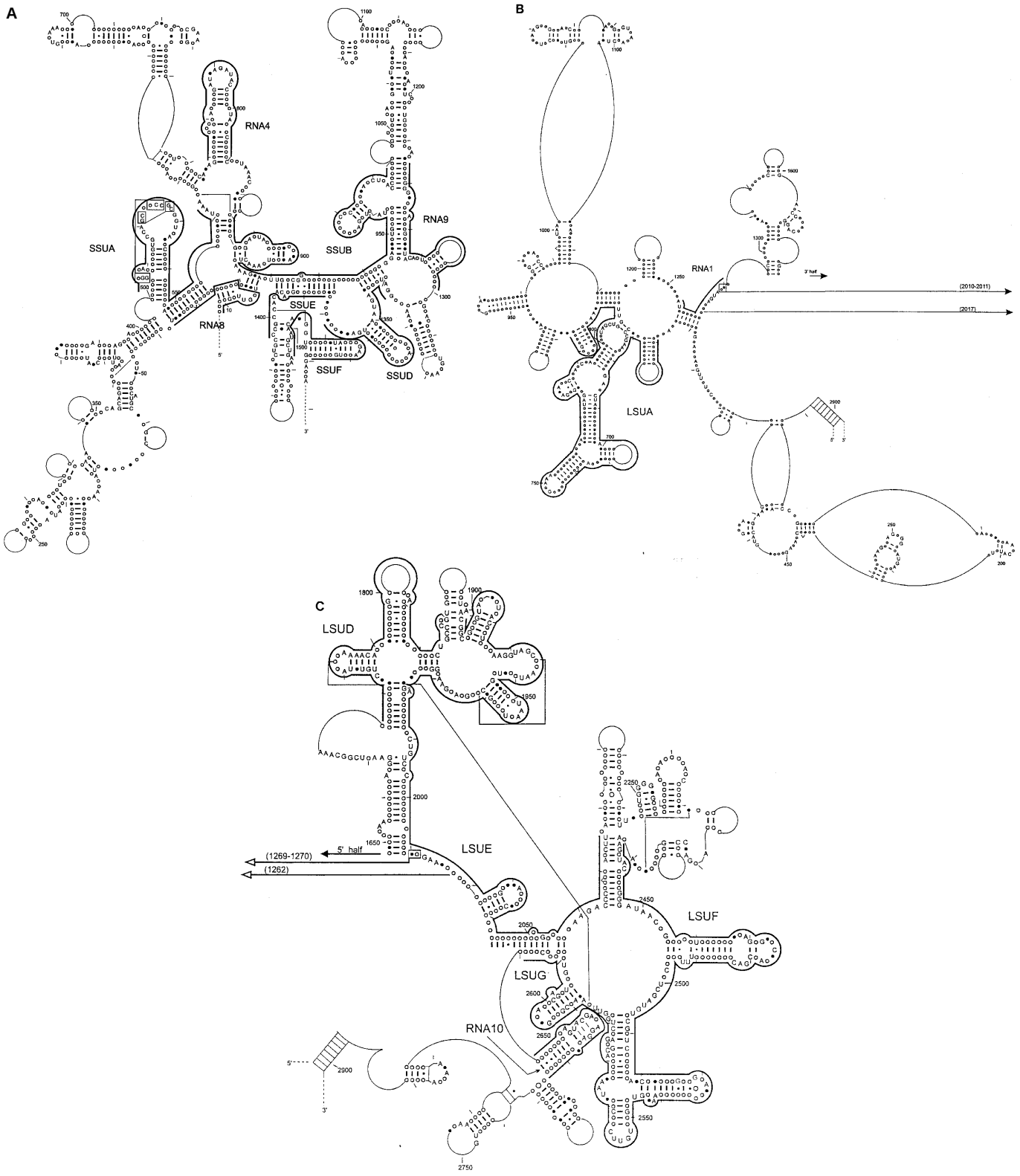

Figure 4. Consensus secondary structure for SSU and LSU rRNAs. Consensus secondary structure figures, based on comparisons of sequences from three phylogenetic domains and two organelle types, for SSU rRNA (**A**), and the 5′ (**B**) and 3′ (**C**) regions of LSU rRNA are shown. They are generally based on *E.coli* structures and are adapted from R. Gutell (26,27). For nucleotide sequence conservation, positions conserved >90% are shown as letters and those conserved 80–90% as filled circles. Open circles identify *E.coli* sites at which 95% of rRNAs examined have a nucleotide. Thin, straight lines indicate long distance interactions and arcs show regions with variable sequence and structure. Areas corresponding to RNAs from the *P.falciparum* 6 kb element are designated by thick lines. Every 10th position, based on *E.coli*, has a tickmark.

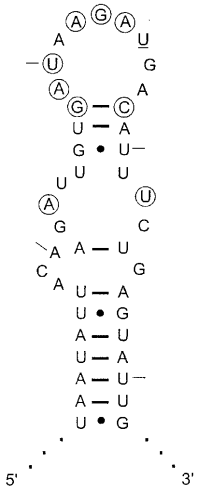

Figure 5. Potential secondary structure for RNA4. The potential secondary structure, based on the three phylogenetic domain, two organelle consensus structure for the 790 loop, of RNA4 is shown. *P.falciparum* positions which match the >95% conserved sequence consensus are circled and those which match the 90–95% consensus are underlined. Tickmarks are shown every 10th position, starting at the 5′ end.

structure of that RNA (5), which had an unexpected feature, discussed below. While mitochondrial rRNAs often have idiosyncratic characteristics (29), it was also possible that this difference, and our inability to find a complementary RNA, meant that the proposed secondary structure was flawed. We therefore reconsidered the secondary structure of SSUA RNA.

Much of the SSUA RNA corresponds to the 530 loop of SSU rRNA. This structure normally has a terminal loop and a six nucleotide bulged loop on the left hand side of its stem. Three nucleotides of the bulged loop interact in a functional pseudoknot with three nucleotides from the terminal loop (30). The previously proposed secondary structure for the *P.falciparum* SSUA RNA lacked the expected left hand bulged loop and instead had one on the right side of the stem (5). After re-evaluation, our newly proposed secondary structure for this region of SSUA RNA (Fig. 6A) lacks the right side loop and has gained a six nucleotide left side loop. The number of basepaired positions in the helix below the bulge is three, less than the five pairs usually seen. The resulting SSUA structure now resembles that of the consensus 530 loop more closely but the potential for a three pair pseudoknot is still not present. Based on positions that are normally involved in pseudoknot formation, the GAU from the bulged loop would be expected to pair with ACA in the terminal loop (Fig. 6A). This predicts Watson–Crick pairing only between the U at position 6 and the A at position 23, rather than all three potential pairs. However, another highly conserved interaction (31) within the terminal loop itself is maintained (Fig. 6A).

The revised SSUA RNA structure places a different sequence at the sites expected to complement the 5′ end sequence. Two sections of sequence from RNA8 can complement the SSUA sequence. Each shows different aspects of positional and sequence conservation. The consensus region corresponding to the SSUA/RNA8 helix has 10 basepairs (bp) and a bulged nucleotide on the RNA8 side (Fig. 4). One alignment of RNA8 with SSUA (Fig. 6A) produces a helix that can basepair at six of the 10 sites. Two of the remaining sites would be G·A pairs. Although these are not certain interactions, G·A pairs are the most common non-canonical interactions (32). In contrast, the alternate possible interaction (Fig. 6B) has only an eight member SSUA/RNA8 helix. However, seven of the eight pairs show conventional basepairing and there is a positionally conserved bulged nucleotide

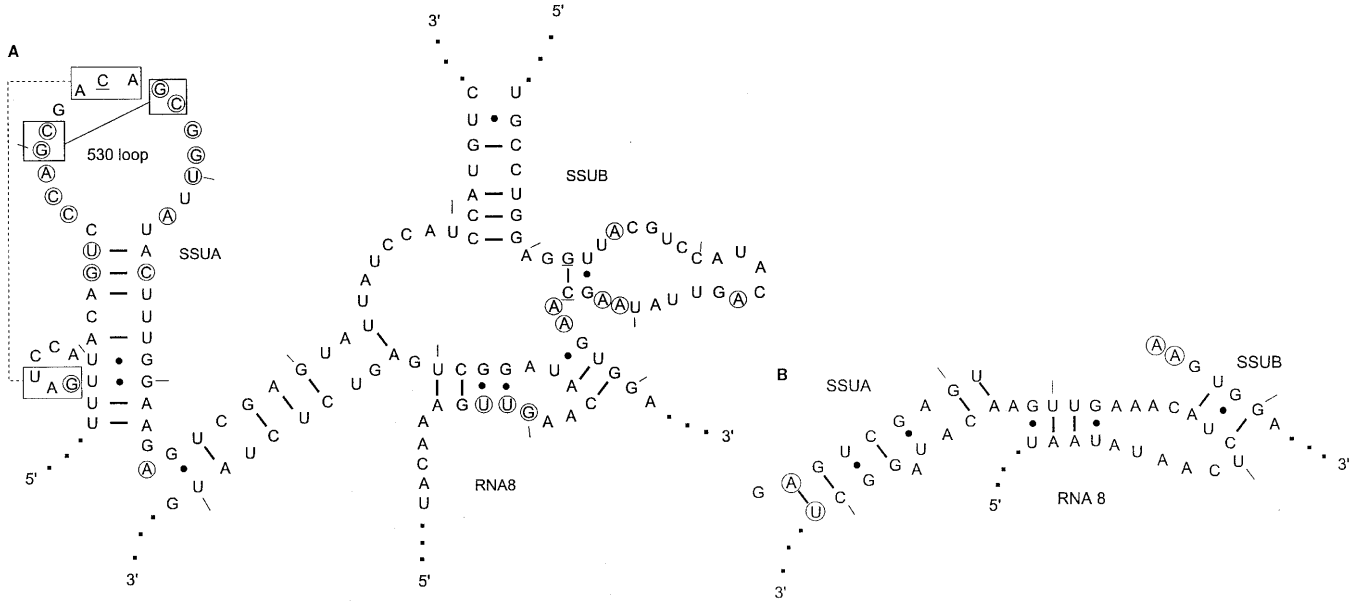

Figure 6. Identification of possible interactions for RNA8. The revised secondary structure, based on consensus SSU rRNA structure, of SSUA RNA is shown, with its potential interaction with RNA8 (**A**). Potential for a short helix between RNA8 and SSUB RNA and for the 5′ end of RNA8 with itself are also indicated. An alternate structure for RNA8 interactions is also shown (**B**), with the SSUA and SSUB structures shown in less detail (starting at SSUA 43 and SSUB 35 in panel A). RNA8 position 1 in (A) corresponds to position 12 in (B). Positions involved in predicted tertiary interactions are boxed. Dotted lines indicate the expected pseudoknot interaction discussed in the text. Other symbols are as described in Figure 5.

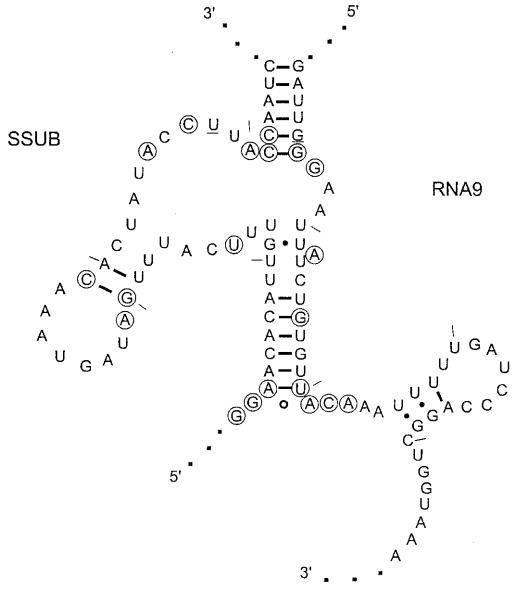

Figure 7. Potential secondary structure for RNA9. The potential secondary structure, based on the SSU rRNA consensus structure, for RNA9 is shown, with its proposed interaction with the 3′ portion of the SSUB RNA. Other symbols are as described in Figure 5.

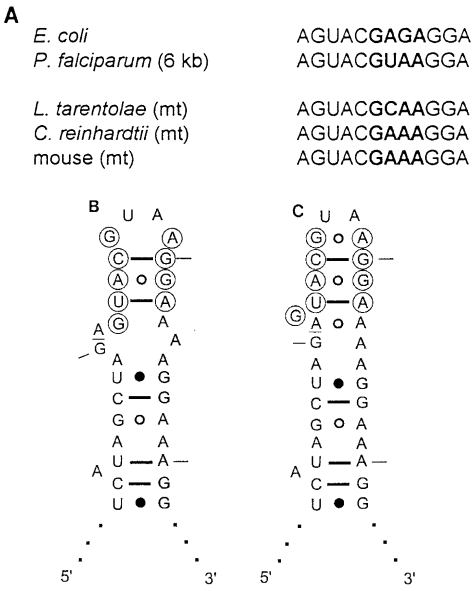

Figure 8. Sequence comparison and potential secondary structure of the sarcin/ricin loop. The 12 nt highly conserved sequences from the sarcin/ricin loop from *E.coli*, several mitochondrial rRNAs, and the corresponding sequence from the *P.falciparum* 6 kb element are aligned (**A**). The GNRA tetraloop sequence is shown in boldface type. (**B**) shows the proposed secondary structure for RNA10, based on consensus figures and (**C**) shows an another possibility, based on detailed structural analysis of a synthetic oligonucleotide related to the rat sequence for this region (34). Other symbols are as described in Figure 5.

(position 27) which is absent from the first structure. In addition, in the three domain, chloroplast consensus (27), nucleotides corresponding to positions 44 and 45 in SSUA are AG in >95% of rRNAs examined. Similarly, the >95% consensus sequence for positions 29–31 of RNA8 (Fig. 6B) is GCU in that comparison. When mitochondrial SSU rRNAs are added to the consensus, three of the five sites (45 in SSUA and 29,30 in RNA8) are reduced to 80–90% sequence conservation (Fig. 4A). The first structure (Fig. 6A) conserves the AG in SSUA but none of the sites in RNA8, although basepairing is preserved for all three, presuming G·A basepairing between positions SSUA 44 and RNA8 31. The alternate structure (Fig. 6B), on the other hand, conserves all five positions.

RNA8 is also expected, from consensus structures, to make a 3 bp helix with part of SSUB and a 4 bp helix with itself. Both of these are possible with both suggested structures but the positional conservation is better with the first structure (Fig. 6A). For RNA8, there should normally be four unpaired nucleotides 5′ of the SSUB/RNA8 helix and two following it. In the first alignment of RNA8 (Fig. 6A), the size of those regions is reduced to 3 and 1 nt, respectively. In the alternate alignment (Fig. 6B), the corresponding regions are 6 and 4 nt. Further, the three domain, two organelle SSU consensus includes a UUG (positions 8–10 in Fig. 6A) that is conserved in >95% of SSU rRNAs. This is missing from the alternate structure (Fig. 6B) but is present in approximately the right place in the first structure (Fig. 6A). The major difference between the RNA8 and consensus SSU rRNA structures in this case is that the UUG is unpaired in the consensus structure but both Us are paired in RNA8. There are thus points for and against each of the possible structures; at the present time, it is not possible to choose between them with confidence.

Comparisons with other SSU rRNA secondary structures suggested that the 3′ end of SSUB RNA should form a helix with a transcript from the SSUC rDNA region (5). However, efforts to detect an SSUC RNA, by RNA blotting with a variety of probes and by RNase protection, were unsuccessful. We have now found that RNA9, a 53 nt transcript, has sequences which are complementary to the SSUB sequences expected to participate in the helix (Fig. 7). It exhibits both sequence and positional conservation. Sequences match a >90% consensus at positions 6–8, 12, 16 and 20–23 (Fig. 7). Sequences conserved in the SSUB portion of the helix normally include a UG at positions 7 and 8. In SSUB, the corresponding nucleotides are CA, preserving the basepairing, though not the specific sequence. RNA9 also has potential to form a short stem loop toward its 3′ end, corresponding to a positionally conserved stem with a non-conserved terminal structure (27).

New LSU rRNA fragments

In addition to the RNAs similar to SSU rRNA, two of the newly described RNAs have been identified as probable components of the fragmented LSU rRNA. One of these corresponds to the sarcin/ricin loop. This sequence is found near the 3′ end of the LSU rRNA. Its central 12 nt are almost universally conserved and it is sensitive to site-specific endonucleolytic cleavage by α-sarcin and to depurination by ricin. Treatment with either of these toxins abrogates the elongation step of protein synthesis and thus signifies a crucial role for this region of rRNA (33). The structure of the sarcin/ricin loop has been extensively studied and has a GAGA tetraloop, which contains the site of ricin and α-sarcin action, at the tip of the stem (34). In mitochondrial LSU rRNAs, the tetraloop sequence is often altered (Fig. 8A) but still

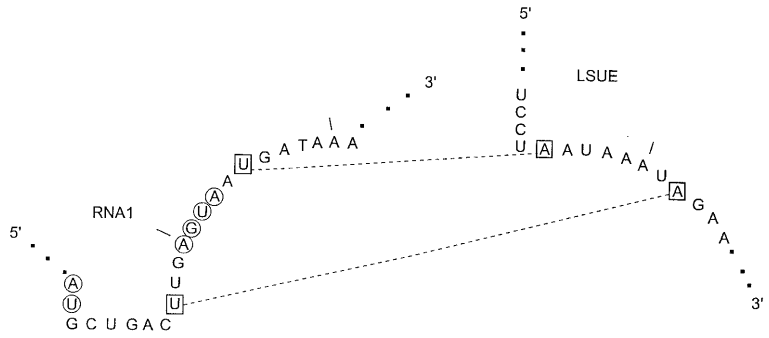

Figure 9. Potential secondary structure for RNA1. The potential secondary structure for RNA1 is shown, with its proposed interaction with the LSUE RNA. Sites involved in the tertiary interactions are boxed. Other symbols are as described in Figure 5.

fits a motif very common in rRNAs: GNRA, where N is any nucleotide and R is a purine (35). The 6 kb element contains a sequence which matches the terminal consensus sequence exactly except for the tetraloop region, which bears the sequence GUAA (Fig. 8A) and is thus consistent with the common tetraloop motif. An oligonucleotide complementary to the putative sarcin/ricin loop sequence described here detects the 100 nt RNA10 (Fig. 3, panel 10). The sequences flanking the sarcin/ricin loop region form an extended helix (Fig. 8B) which corresponds to a positionally conserved helix in the consensus structure (Fig. 4). Detailed studies of the conformation of a synthetic oligonucleotide based on the rat sarcin/ricin loop suggest that the G at position 12 is bulged and the As at positions 11 and 23 form a non-canonical pair (34). This structure is also possible with the *P.falciparum* sequence (Fig. 8C).

The bulged A at position 3 of the sarcin/ricin loop (Fig. 8B and C) is not expected from consensus diagrams. Its location is suggested by comparison with the corresponding *Theileria parva* sequence (36). This related parasite also has fragmented mitochondrial rRNAs. Its sarcin/ricin loop sequence matches that of *P.falciparum* exactly, except for three positions (8, 24 and 32) at which the changes retain potential basepairing and for the lack of a nucleotide corresponding to the bulged A. The absence of the bulged A in *T.parva* and the improved basepairing of positions 1, 2 with 32, 31 afforded by bulging position 3 support this departure from overall consensus structures.

The second RNA with similarity to LSU rRNA is the 96 nt RNA1. It contains sequence which matches 20 of 23 nt of sequence from a central region of *E.coli* LSU rRNA (Fig. 9). Six of these are conserved at >95% in the three domain, two organelle consensus (26). The highly conserved 5′ AU (positions 1, 2) is normally separated from a 15 nt single-stranded region by one side of a six member helix, formed with sequences further 5′. The expected half helix is positionally conserved in RNA1, though the 6 kb element sequence which provides the other half of the helix is not yet identified. Of the 15 nt single-stranded region, 3 nt are normally involved in a tertiary interaction with a portion of LSU rRNA (26) that in *P.falciparum* is contained within LSUE RNA (5). Two of the three interactions are possible in the *P.falciparum* sequences (Fig. 9). The third interaction expected (position 14 in RNA1 with position 6 in LSUE in Figure 9) does not maintain a canonical pairing, as both nucleotides are As. Such non-canonical pairings are known (31) but there is insufficient information regarding the precise arrangement and interactions of these unusual rRNAs to conclude that it occurs in this case.

Two of the originally described rDNA regions, LSUB and LSUC, correspond to the GTPase center of the LSU rRNA. As with the SSUC rDNA region, RNA blotting and RNase protection experiments have not detected abundant transcripts for these rDNA regions. It is possible that one or more of the remaining five new RNAs might function as the GTPase center, similar to the replacement of SSUC with RNA4. However, a search for sequences corresponding to the GTPase center found no convincing matches from the regions encoding RNA2, RNA3, RNA5, RNA6 or RNA7, nor have matches to any other conserved rRNA regions yet been found for these RNAs. Their small size and abundance suggest that they are also fragments of rRNA but their roles remain to be determined.

DISCUSSION

The RNA blotting experiments presented here have demonstrated abundant transcripts for 10 of the originally proposed rDNA regions of the *P.falciparum* 6 kb element. This set of RNAs corresponds to regions critical for ribosome function but is missing important sequences expected in an organelle rRNA. Transcript mapping has now detected 10 additional small transcripts. For five of these, specific locations in secondary structure can be predicted, filling in major gaps from the previously proposed structures. With the exception of the GTPase center, the earlier identification of which is now in question (see below), by far the majority of regions with highly conserved nucleotide sequence expected in rRNAs are now represented. Identification of rRNAs corresponding to structures that are predominantly positionally conserved is more difficult and remains an issue for future study. The identification of these new rRNA fragments further supports the potential for the small transcripts from the 6 kb element to participate in functional ribosomes.

The newly identified rRNA fragments generally include shorter contiguous regions of strong similarity than the originally identified rDNA sequences, explaining the difficulty in finding them. Based on the functions and/or conservation of corresponding regions of other rRNAs, however, they constitute important portions of the rRNA and presumably the mitochondrial ribosomes. The 790 loop (RNA4) is implicated in the interaction of the large and small subunits during protein synthesis and alteration of its

sequence or ability to hydrogen bond decreases subunit association (reviewed in 37). The abrogation of translation by the action of the ribotoxins α-sarcin and ricin (33) attest to the importance of the sarcin/ricin loop (RNA10). The identities of RNA8 and RNA9 rest to a significant degree on their ability to appropriately complement sequences of other rRNA fragments, and are bolstered by occurrence of a few highly conserved nucleotides and positional conservation of their potential structure. These rRNA regions have less clearly defined roles than the 790 loop and sarcin/ricin loop but correspond to generally well-conserved regions. RNA1 has maintained sufficient similarity to *E.coli* rRNA for its identification, has half a dozen highly conserved nucleotides, and has maintained two of three predicted tertiary interactions with LSUE. Experiments to verify the association of these and other rRNA fragments into ribosomes are in progress.

The combination of the discontinuous and scrambled nature of the 6 kb element rDNA sequences and the unconventional secondary structures often found for small mitochondrial rRNAs makes it particularly difficult to determine the identity of each of the small transcripts. No rRNA similarities have yet been found for five of these although their size and abundance make them excellent candidates for rRNA fragments. While further sequence analysis may reveal hints to their identities, it is also possible that they correspond to positionally conserved regions that are not well-conserved in sequence. Alternately, they may exhibit unique structures called for by the fragmented nature of the *P.falciparum* mitochondrial rRNAs. They may, for example, interact with other fragments to maintain a particular secondary structure and bring important nucleotides into the proper orientation and proximity. Presumably such interactions between RNAs and/or with proteins in ribosomes allow the various rRNA fragments to maintain the correct spatial arrangement for function. In such a situation, there may be important structural roles for (some of) the five as yet unidentified small RNAs which do not require extensive sequence or positional conservation but only the ability to interact with other rRNAs at selected sites. It is also possible that some of them have undiscovered functions unrelated to rRNA.

Abundant, appropriately sized transcripts were not detected for three of the originally proposed rDNA regions, SSUC, LSUB and LSUC, by RNA blotting or RNase protection assays. These regions are transcribed, as demonstrated by data showing extensive polycistronic transcription of the *P.falciparum* 6 kb element (38). The polycistronic transcripts are not very abundant, suggesting rapid processing to the smaller RNAs, and those portions of transcript which correspond to the 'missing' RNAs may be relatively rapidly turned over. The lack of easily detectable, abundant transcripts suggests that the earlier identification of the *P.falciparum* SSUC, LSUB and LSUC rDNA regions was in error.

The absence of a transcript corresponding to SSUC is not critical since the abundant small transcript designated RNA9 has the potential to fill the same position in the secondary structure, and presumably the same function in the ribosome. The absence of LSUB and LSUC RNAs is more troubling since they were predicted to comprise the GTPase center, which is required for protein synthesis. Three possibilities may explain this apparent absence. The GTPase center functions may be provided by one or more of the remaining unidentified small transcripts from the 6 kb element, RNAs with the GTPase sequences may be imported from outside the mitochondrion, or the LSUB and LSUC sequences may be transcribed but then modified in such a way as to render them unable to hybridize to the probes we employed. The first of these possibilities may require a novel sequence or structure to provide the GTPase function, since no significant similarity has yet been detected between the unidentified RNAs and conserved GTPase sequences. As to the other possibilities, small RNAs are known to be imported into mitochondria (39,40), and post-transcriptional modification of rRNA sequences has also been reported (41–43). Sorting out these possibilities will probably require analysis of the RNA components in the mitochondrial ribosomes.

As with the GTPase center, the deviations of SSUA RNA sequence from the generally conserved characteristics of the 530 loop provoke questions. The 530 loop has been extensively studied and specific sites in and around it have been implicated in interactions with EF-Tu (44), polysome formation and/or stability (45), and interactions with ribosomal proteins S12 and S4, which are associated with the fidelity of translation (46,47). Some of the critical sites are maintained in SSUA RNA, notably a G in the site corresponding to *E.coli* position 530 (position 29 in SSUA, Fig. 6A). In contrast, studies with *E.coli* mutations suggest that substitution of the G at position 517 with a C increases misreading (48); the corresponding site in the *P.falciparum* SSUA RNA sequence (position 16 in SSUA, Fig. 6A) is a C. Perhaps most importantly, potential for the pseudoknot expected between the bulged loop and terminal loop appears decreased, with only one of the set of three basepairs able to form a canonical pair. Mutational analysis has shown that the absence of basepairing potential between these sites results in decreased cell viability and even substitution of wobble basepairing for Watson–Crick interactions has negative functional consequences (30). However, there are a few examples of mitochondrial rRNAs which do not exhibit strong basepairing potential, such as those from kinetoplastid protozoa (49–51). The lack of expected basepairing for the 530 loop pseudoknot may thus still allow functioning ribosomes.

The characteristics of the small abundant RNAs are, for the most part, consistent with a functional protein synthetic apparatus, albeit probably a most unusual one. The loss of sequence, compared with eubacterial rRNAs, may reflect adaptation to the very small size of the mitochondrial genome in *Plasmodium*, with less critical sequences being evolutionarily jettisoned. *Theileria parva*, a related parasite, also contains a very small mitochondrial genome with highly fragmented rRNAs (36), and these unusual rRNAs may be characteristic of other members of phylum Apicomplexa as well (reviewed in 52). Evidence for the existence of the protein products of the mitochondrial mRNAs (reviewed in 52) suggests that the *Plasmodium* mitochondrion has developed mechanisms for overcoming the challenges posed by the extreme fragmentation, small fragment size, and apparent absence of some sequences and expected interactions. Deciphering them may provide insights into the role of RNA structure in ribosome function.

ACKNOWLEDGEMENTS

We thank Mark Drew, Holli Banister and Julie Anderson for technical assistance, Shaofeng Yan for work performed during a graduate student rotation, and Martha Thorning for invaluable assistance with figure preparation. We further thank Dr Ira G. Wool for helpful discussions about the sarcin/ricin loop, Drs Robin Gutell, Michael Gray and Murray Schnare for suggestions on secondary structures, and Drs Gerard Cangelosi and Marilyn Parsons for comments on the manuscript. This work was

supported by NSF grant MCB 9205809 to J.E.F., who is also a Burroughs Wellcome New Investigator in Molecular Parasitology.

REFERENCES

1 Feagin,J.E. (1992) *Mol. Biochem. Parasitol.* **52**, 145–148.
2 Vaidya,A.B., Akella,R. and Suplick,K. (1989) *Mol. Biochem. Parasitol.* **35**, 97–108.
3 Aldritt,S.M., Joseph,J.T. and Wirth,D.F. (1989) *Mol. Cell. Biol.* **9**, 3614–3620.
4 Gray,M.W. (1989) *Annu. Rev. Cell Biol.* **5**, 25–50.
5 Feagin,J.E., Werner,E., Gardner,M.J., Williamson,D.H. and Wilson,R.J.M. (1992) *Nucleic Acids Res.* **20**, 879–887.
6 Suplick,K., Morrisey,J. and Vaidya,A.B. (1990) *Mol. Cell. Biol.* **10**, 6381–6388.
7 Spencer,D.F., Collings,J.C., Schnare,M.N. and Gray,M.W. (1987) *EMBO J.* **6**, 1063–1071.
8 Schnare,M.N. and Gray,M.W. (1990) *J. Mol. Biol.* **215**, 73–83.
9 Burgin,A.B., Parodos,K., Lane,D.J. and Pace,N.R. (1990) *Cell* **60**, 405–414.
10 Skurnick,M. and Toivanen,P. (1991) *Mol. Microbiol.* **5**, 585–593.
11 Heinonen,T.Y.K., Schnare,M.N., Young,P.G. and Gray,M.W. (1987) *J. Biol. Chem.* **262**, 2879–2887.
12 Boer,P. and Gray,M.W. (1988) *Cell* **55**, 399–411.
13 Lemieux,C., Boulanger,J., Otis,C. and Turmel,M. (1989) *Nucleic Acids Res.* **17**, 7997.
14 Turmel,M., Boulanger,J., Schnare,M.N. and Gray,M.W. (1991) *J. Mol. Biol.* **218**, 293–311.
15 Coté,V., Mercier,J.P., Lemieux,C. and Turmel,M. (1993) *Gene* **129**, 69–76.
16 Schnare,M.N., Heinonen,T.Y.K., Young,P.G. and Gray,M.W. (1986) *J. Biol. Chem.* **261**, 5187–5193.
17 Campbell,D.A., Kubo,K., Clark,C.G. and Boothroyd,J.C. (1987) *J. Mol. Biol.* **196**, 113–124.
18 White,T.C., Rudenko,G. and Borst,P. (1986) *Nucleic Acids Res.* **14**, 9471–9489.
19 Gray,M.W. and Schnare,M.N. (1990) In Hill,W.E., Dahlberg,A., Garrett,R.A., Moore,P.B., Schlessinger,D. and Warner,J.R. (eds) *The Ribosome: Structure, Function, & Evolution.* American Society for Microbiology, Washington, D.C. pp. 589–597.
20 Denovan-Wright,E.M. and Lee,R.W. (1994) *J. Mol. Biol.* **241**, 298–311.
21 Joseph,J.T., Aldritt,S.M., Unnasch,T., Puijalon,O. and Wirth,D.F. (1989) *Mol. Cell. Biol.* **9**, 3621–3629.
22 Feagin,J.E. and Drew,M.E. (1995) *Exp. Parasitol.* **80**, 430–440.
23 Trager,W. and Jensen,J.B. (1978) *Nature* **273**, 621–622.
24 Chomczynski,P. and Sacchi,N. (1987) *Anal. Biochem.* **162**, 156–159.
25 Gardner,M.J., Feagin,J.E., Moore,D.J., Spencer,D.F., Gray,M.W., Williamson,D.H. and Wilson,R.J.M. (1991) *Mol. Biochem. Parasitol.* **48**, 77–88.
26 Gutell,R.R., Gray,M.W. and Schnare,M.N. (1993) *Nucleic Acids Res.* **21**, 3055–3074.
27 Gutell,R.R. (1994) *Nucleic Acids Res.* **22**, 3502–3507.
28 Gardner,M.J., Feagin,J.E., Moore,D.J., Rangachari,K., Williamson,D.H. and Wilson,R.J.M. (1993) *Nucleic Acids Res.* **21**, 1067–1071.
29 Raué,H.A., Klootwijk,J. and Musters,W. (1988) *Prog. Biophys. Mol. Biol.* **51**, 77–129.
30 Powers,T. and Noller,H.F. (1991) *EMBO J.* **10**, 2203–2214.
31 Gutell,R.R., Larsen,N. and Woese,C.R. (1994) *Microbiol. Rev.* **58**, 10–26.
32 Gautheret,D., Konings,D. and Gutell,R.R. (1994) *J. Mol. Biol.* **242**, 1–8.
33 Wool,I.G., Glück,A. and Endo,Y. (1992) *Trends Biochem. Sci.* **17**, 266–269.
34 Szewczak,A.A., Moore,P.B., Chan,Y.-L. and Wool,I.G. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 9581–9585.
35 Uhlenbeck,O.C. (1990) *Nature* **346**, 613–614.
36 Kairo,A., Fairlamb,A.H., Gobright,E. and Nene,V. (1994) *EMBO J.* **13**, 898–905.
37 Dahlberg,A.E. (1989) *Cell* **57**, 525–529.
38 Ji,Y., Mericle,B.L., Rehkopf,D.H., Anderson,J.D. and Feagin,J.E. (1996) *Mol. Biochem. Parasitol.* **81**, 211–223.
39 Simpson,A.M., Suyama,Y., Dewes,H., Campbell,D.A. and Simpson,L. (1989) *Nucleic Acids Res.* **17**, 5427–5445.
40 Chang,D.D. and Clayton,D.A. (1989) *Cell* **56**, 131–139.
41 Lane,B.G., Ofengand,J. and Gray,M.W. (1995) *Biochimie* **77**, 7–15.
42 Bakin,A., Lane,B.G. and Ofengand,J. (1994) *Biochemistry* **33**, 13475–13483.
43 Mahendran,R., Spottswood,M.S., Ghate,A., Ling,M.-I., Jeng,K. and Miller,D.L. (1994) *EMBO J.* **13**, 232–240.
44 Powers,T. and Noller,H.F. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 1364–1368.
45 Powers,T. and Noller,H.F. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 1042–1046.
46 Powers,T. and Noller,H.F. (1994) *J. Mol. Biol.* **235**, 156–172.
47 Alksne,L.E., Anthony,R.A., Liebman,S.W. and Warner,J.R. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 9538–9541.
48 O'Connor,M., Göringer,H.U. and Dahlberg,A.E. (1992) *Nucleic Acids Res.* **20**, 4221–4227.
49 de la Cruz,V.F., Lake,J.A., Simpson,A.M. and Simpson,L. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 1401–1405.
50 de la Cruz,V.F., Simpson,A.M., Lake,J.A. and Simpson,L. (1985) *Nucleic Acids Res.* **13**, 2337–2356.
51 Sloof,P., Van den Burg,J., Voogd,A., Benne,R., Agostinelli,M., Borst,P., Gutell,R. and Noller,H. (1985) *Nucleic Acids Res.* **13**, 4171–4190.
52 Feagin,J.E. (1994) *Annu. Rev. Microbiol.* **48**, 81–104.