Nucleic Acids Research, Vol. 20, No. 4 879–887

Homologies between the contiguous and fragmented rRNAs of the two *Plasmodium falciparum* extrachromosomal DNAs are limited to core sequences

Jean E.Feagin, Erica Werner, Malcolm J.Gardner[1], Donald H.Williamson[1] and Robert J.M.Wilson[1]

Seattle Biomedical Research Institute, 4 Nickerson St, Seattle, WA 98109-1651 and [1]Division of Parasitology, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK

Received September 5, 1991; Revised and Accepted January 13, 1992 GenBank accession no. M76611

ABSTRACT

***Plasmodium falciparum* contains two extrachromosomal DNAs, a 6 kb linear element and a 35 kb circular DNA; both encode rDNA sequences. The 6 kb element rDNAs comprise fragments of both large and small subunit rRNAs. Comparison of these with corresponding rDNA sequences from the 35 kb DNA and *E. coli* show that sequences conserved between the three are largely confined to highly conserved core regions; in fact, most of the 6 kb rDNA sequences correspond to core regions. Both the 6 kb element and 35 kb rDNAs show less conservation to each other than to *E. coli* sequences, suggesting that the two extrachromosomal DNAs of *P. falciparum* are not closely related. The characteristics of the fragmented rRNAs from the 6 kb element suggest they are functional, possibly in mitochondrial ribosomes.**

INTRODUCTION

Two extrachromosomal DNAs with organelle-like characteristics have been found in malarial parasites. One is a 6 kb, tandemly repeated element which encodes mitochondrial protein coding genes and fragments of large and small rRNA genes (1–4). It is likely that this DNA comprises at least part of the mitochondrial genome since cytochrome oxidase I and cytochrome *b*, universally encoded by mitochondrial genomes (5), are encoded by it. Its small size would make it by far the smallest mitochondrial genome yet described. The molecule also encodes fragments of rDNA sequence. The second extrachromosomal DNA is a 35 kb circular molecule which encodes, in the region thus far sequenced (approximately half the molecule), complete large and small rRNAs, tRNAs, two subunits of RNA polymerase, and a ribosomal protein gene (6–10). Unexpectedly, some characteristics of the 35 kb DNA are reminiscent of chloroplast genomes, including the arrangement of the rRNA genes, which are duplicated in an inverted repeat. Moreover, the protein coding genes identified to date are more commonly found in chloroplast than mitochondrial genomes.

Malarial parasites thus appear to have unusual organelle DNAs; although neither extrachromosomal DNA has been unambiguously located subcellularly, they may be located in different organelles (9). The small size of the 6 kb element suggests that if it is the sole mitochondrial genome, a large proportion of the usual mitochondrial coding functions have been lost or reside elsewhere, either transferred to the nucleus as reported for other organisms (11,12) or present on the 35 kb DNA. If the latter is true, some degree of cooperation between the two genomes may be required to produce a functional mitochondrion. This would also suggest that the extrachromosomal DNAs might have originated from a common progenitor. Alternatively, the 35 kb DNA may have an unrelated and as yet unknown function.

Sequence data for the 6 kb element is available from *P. yoelii* (1) and *P. gallinaceum* (2,4) which are murine and avian parasites, respectively, whereas the 35 kb sequence is from the human parasite *P. falciparum* (6,7,10) Wilson *et al.*, in preparation). We have determined the sequence of the 6 kb element from *P. falciparum* to allow intraspecific comparison of sequences from the two extrachromosomal DNAs. Such comparisons allow assessment of the degree of relatedness between the two molecules and may provide insight into the possible interrelationships between them. We have found that cross-hybridization between the molecules is minimal and probably due to rDNA sequences. Comparative analyses of the rDNA sequences of the 6 kb element and 35 kb circular DNA of *P. falciparum* indicate that homology is localized to highly conserved core sequences; indeed, the 6 kb element rDNA fragments are predominantly composed of core sequences. In these regions, the two *P. falciparum* DNAs are more homologous to *E. coli* than to each other. These findings suggest that the 6 kb element and the 35 kb circular DNA are not closely related. The characteristics of the fragmented rRNAs from the 6 kb element suggest that they function in mitochondrial ribosomes.

METHODS

Parasites

The C10 clone of *P. falciparum* (13) was used for these studies. Parasites were cultivated by the method of Trager and Jensen (14). Intact parasites were prepared by lysis of infected

erythrocytes with saponin or 1% acetic acid (15), followed by washes with phosphate-buffered saline. Isolated parasites were used immediately or quick-frozen for later use.

Nucleic acid isolation

Total parasite DNA was prepared by proteinase K digestion and RNase treatment (16). *P. knowlesi* 35 kb circular DNA was isolated from CsCl gradients, as previously described (6). Total RNA was prepared from a mixed population of parasites (approximately 50% trophozoites) by the acid-guanidinium-phenol-chloroform method (17) as previously described (10).

Isolation, cloning, and sequencing of the 6 kb element

Total *P. falciparum* DNA was isolated and fractionated on CsCl/DAPI gradients by modification of the method developed for isolation of the 35 kb circular molecules (6). Briefly, infected erythrocytes were extracted with sodium lauroyl sarcosine in the presence of proteinase K, and the extract was adjusted to 31.2 ml with sterile saline (0.1 M NaCl), mixed with 29.6 gm solid CsCl and 0.8 ml of 10 mg/ml DAPI, and centrifuged at 45,000 rpm for 24 hr in a Beckman VTi50 rotor. At this stage, the 6 kb DNA forms a diffuse, weakly fluorescent band well below the main chromosomal band. Subsequent purification of the 6 kb DNA involved two sequential centrifugations in 5 ml gradients in a Beckman VTi65 rotor; the first centrifugation was for 4 hr at 65,000 rpm and the second for 12 hr at 33,000 rpm. For these runs, the samples were topped up with a CsCl solution (4.3 gm CsCl plus 4 ml sterile saline). Additional DAPI was either omitted or added at 1−2 μl per gradient, as adding too much DAPI at this stage often leads to irreversible precipitation of a DAPI/DNA complex. The 6 kb molecules formed a diffuse weakly fluorescent band about 1 cm wide, approximately 0.5 cm below any remaining band of chromosomal DNA. Following removal by side puncture, the DNA was purified by phenol extraction and ethanol precipitation, digested with *Hin*dIII, and the three resulting fragments were cloned into pBluescript (Stratagene) for sequencing. These clones were sequenced by the dideoxy chain termination method using Sequenase (U.S. Biochemical), according to manufacturer's instructions. To verify the junctions between the *Hin*dIII fragments, PCR products overlapping the *Hin*dIII sites were cloned and sequenced. Virtually all positions were confirmed by sequencing both strands and all positions were sequenced multiple times.

Blotting and hybridization

For Southern blots, total *P. falciparum* DNA digested with restriction enzymes was electrophoresed on 0.7% agarose gels in TAE (40 mM Tris-acetate, 1 mM EDTA) and transferred to Nytran (Schleicher and Schuell) by capillary blotting. Probes were prepared by the random priming method (18) and hybridizations and washes were performed as described elsewhere (10). For RNA blots, total *P. falciparum* RNA was electrophoresed on 12% acrylamide, 7 M urea gels in TBE (0.1 M Tris-borate, 0.9 mM EDTA) and electrophoretically transferred to Nytran in TAE. Probes for RNA blots were a ^{32}P-labeled *in vitro* antisense transcript complementary to the gene and some flanking sequence or a ^{32}P-endlabeled oligonucleotide (GGAGTCTCACACTA-GCGACAATGGGG). For the *in vitro* transcript probe, the blot was prehybridized overnight at 42°C in 5×SSPE (1×SSPE: 90 mM NaCl, 10 mM NaH_2PO_4, 1 mM EDTA), 50% deionized formamide, 1% SDS, 150 μg/ml denatured salmon sperm DNA, and 0.1% each Ficoll, polyvinylpyrrolidone, and bovine serum albumin. Hybridization was in the same buffer at 60°C and washes, all at 60°C, were 2×15 min in 2×SSPE, 0.1% SDS and 2×15 min in 0.2% SSPE, 0.1% SDS. Hybridization and wash conditions for the oligonucleotide probe were as previously described (10), at 70°C.

Sequence alignments and potential secondary structures

The rRNA sequences from the 35 kb circular DNA were aligned with the *E. coli* rRNA sequences using the LINEUP program of the UWGCG package (19). rDNA sequences from the 6 kb element were aligned relative to these by eye, based on the predicted linear arrangement of the fragmented rRNAs. Adjustments were made to account for conserved secondary structure. Predicted secondary structures for the 35 kb-encoded rRNAs were taken from Gardner, *et al.* (10) and Wilson *et al.* (in preparation). The predicted 6 kb element rRNA secondary structures are based on those proposed for the *P. gallinaceum* 6 kb element (4), with alterations as required by nucleotide differences between the species.

RESULTS

Cross-hybridization between the 6 kb and 35 kb DNAs

To assess the possibility that the 6 kb and 35 kb extrachromosomal DNAs of malaria parasites are related, duplicate Southern blots of *P. falciparum* genomic DNA were probed with ^{32}P-labeled 35 kb DNA isolated from *P. knowlesi* (20) and the cloned *P. yoelii* 6 kb element (1) (Fig. 1). In addition to several strongly

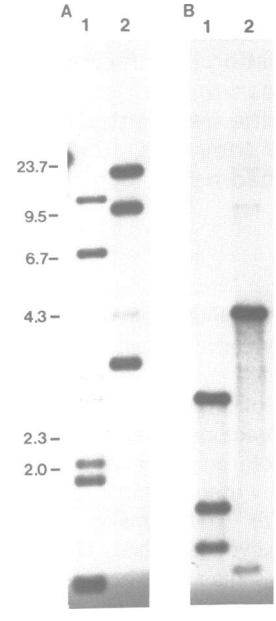

Figure 1. Southern blot analysis of cross-hybridization between the 6 kb element and 35 kb circle. Total *P. falciparum* DNA (2 mg/lane) was digested with restriction enzymes (lane 1, *Hin*dIII; lane 2, *Hpa*I), electrophoresed on an agarose gel, transferred to nylon membrane, and probed with ^{32}P-labeled isolated 35 kb DNA from *P. knowlesi* (A) and cloned 6 kb element from *P. yoelii* (B). The third largest *Hin*dIII fragment detected by the *P. knowlesi* probe (A), although hybridizing less strongly, is a component of the *P. falciparum* 35 kb circular DNA (6); the lesser degree of hybridization may reflect some heterogeneity between *P. falciparum* and *P. knowlesi* in this region. Size markers are indicated in kb.

hybridizing bands, the *P. knowlesi* 35 kb DNA probe detects lightly hybridizing bands which correspond to bands detected strongly with the *P. yoelii* 6 kb element probe. Similarly, the 6 kb element probe hybridizes slightly to some fragments which are detected strongly by the 35 kb DNA probe. As described below, the fragments which are detected with both probes contain rDNA sequences; for example, the most evident 6 kb element-derived fragments detected with the 35 kb DNA probe, a 2.9 kb *Hin*dIII fragment and a 5 kb *Hpa*I fragment, contain the regions with the largest highly conserved rDNA sequences. No other significant cross-hybridization was detected, suggesting that the two genomes are most similar in the rDNA regions.

Sequence of the *P. falciparum* 6 kb element

To further examine the relationship between the 6 kb element and 35 kb circular DNA, we have sequenced the 6 kb element from the same cloned line of *P. falciparum* we previously used in studies of the 35 kb circular DNA. We here report the sequence of the *P. falciparum* 6 kb element (Genbank accession number M76611) and analysis of its rDNA sequences. The element (Fig. 2) contains three open reading frames with homology to cytochrome oxidase I, cytochrome oxidase III, and cytochrome *b*. The nucleotide sequence of the *P. falciparum* 6 kb element is 89% conserved relative to the nucleotide sequence of the corresponding element from *P. yoelii* (1). The open reading frames are approximately 86% conserved between these species and the predicted proteins are 98% composed of conserved or conservatively replaced residues. The 6 kb element appears, therefore, to be very highly conserved between these species.

We have identified regions encoding fragments of both large (lsu6) and small (ssu6) rRNAs (Fig. 2). The 13 rDNA regions discussed here are more highly conserved between *Plasmodium* species at the nucleotide level than are the protein coding genes; all but one are greater than 90% conserved relative to corresponding *P. yoelii* sequences (Table 1). The order of these rDNA sequences in the element is also conserved between species. The 5′ and 3′ ends of the *P. falciparum* rDNA fragments have yet to be experimentally determined and the sequences selected for analysis are therefore based on predicted secondary structure, including only those nucleotides predicted to be in a helical structure. The only exceptions are the 3′ ends of fragments ssu6A and ssu6D and the 5′ ends of lsu6E and lsu6F; these extend 9, 4, 3, and 9 nucleotides from the closest helix and were included because they are conserved in a variety of species. The sizes given for the rDNA fragments in Table 1 thus represent minimum estimates and actual sizes may be significantly larger. Depending on the proximity of other genes, any number from a few to a few hundred additional nucleotides may be included, although preliminary mapping studies suggest smaller additions are more likely. The 5′ end of *P. yoelii* lsu6E transcript has been mapped (21) and is 3 nt further 5′ than shown for *P. falciparum* lsu6E (Fig. 5).

The minimum size estimates, based on the above criteria, for lsu6 and ssu6 fragments range from 28−178 nt. All but two of the fragments are transcribed in the same orientation and many of these are in close proximity, suggesting the possibility that some may be co-transcribed and processed. With the exception of lsu6A and B and lsu6D and E, however, no two fragments predicted to be adjacent to one another in rRNA are adjacent in the DNA sequence of the 6 kb element (Fig. 2). Rather, they are out of order and lsu6 and ssu6 fragments are interspersed between each other and the protein coding genes. The potential 5′ end of lsu6G overlaps the 3′ end of ssu6B by two nucleotides; experimental determination of the ends of transcripts from these fragments will be necessary to assess the actual extent of overlap. The 3′ ends of the CYb open reading frame and fragment lsu6B also have a predicted overlap of 18 nucleotides. As they are transcribed in opposite orientations, this overlap presumably does not affect simultaneous expression of both products.

Table 1. Characteristics of *P. falciparum* 6 kb element-encoded rRNAs.

fragment	min. size[1]	core size[2]	% conservation between[3] Py/Pf	Ec/35/6	Ec/35	Ec/6	35/6
ssuA	86	54	98	51.9	75.9	66.7	55.6
ssuB	116	106	98	50.9	75.5	58.5	61.3
ssuC	75	28	93	10.7	53.6	35.7	14.3
ssuD	53	42	100	54.8	73.8	69.1	54.8
ssuE	28	28	100	73.8	85.7	82.1	73.8
ssuF	71	53	100	47.2	71.7	62.3	54.7
total ssu	429	311	98	49.5	73.6	62.1	55.3
lsuA	161	129	99	51.2	69.0	65.1	63.6
lsuB	92	24	85	54.2	66.7	75.0	58.3
lsuC	94	18	90	66.7	77.8	77.8	66.7
lsuD	97	51	95	52.9	68.6	56.8	68.6
lsuE	178	164	99	62.2	77.4	70.7	68.9
lsuF	100	91	99	56.0	76.9	62.6	63.7
lsuG	100	92	97	63.0	83.7	70.7	68.5
total lsu	822	569	94	57.8	75.2	64.2	63.4

[1]Minimum predicted size, based on secondary structure (see text).
[2]Size of core region encompassed by the minimum predicted fragment.
[3]Total positions conserved relative to size of fragment. Comparison of *P. yoelii* (Py) to *P. falciparum* (Pf) sequences encompasses the entire minimum sized sequences. Comparisons between *E. coli* (Ec), 35 kb DNA-encoded (35), and 6 kb element-encoded (6) rRNAs include only core regions presented in Figs. 4 and 5.

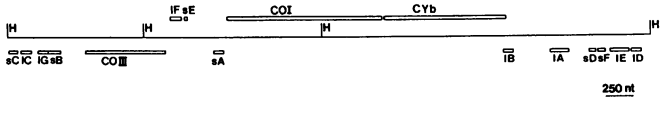

Figure 2. Map of the *P. falciparum* 6 kb element. A schematic map of the element is shown with identified genes indicated above or below the line, according to the direction of transcription (left to right above and right to left below the line). The *Hin*dIII sites (H) are indicated. CYb, cytochrome *b*; COI, cytochrome oxidase I; COIII, open reading frame with some homology to cytochrome oxidase III; lA-G, fragments of large rDNA sequence; sA-F, fragments of small rDNA sequence. The rDNA fragments are named in order of linear occurrence in corresponding complete rRNAs.

rRNA transcripts from the 6 kb element

Transcripts for several of the 6 kb element rDNA fragments have been found in *P. falciparum* (Fig. 3). In some cases, a single transcript of a size consistent with the rDNA fragment being investigated was detected. The ssu6A fragment, for example, has a predicted minimum size of 86 nt (Table 1) and a single 115 nt transcript (Fig. 3A). Similarly, the lsu6E fragment has a predicted size of 178 nt and a single 185 nt transcript (Fig. 3B). In some instances, more than one transcript is detected. The sizes of such transcripts are generally consistent with co-transcription and processing of adjacent rDNA regions but further transcript mapping will be required to assess this possibility.

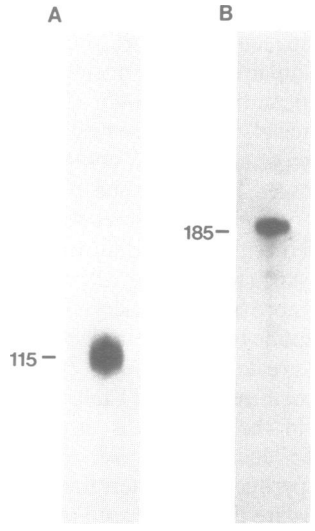

Figure 3. RNA blot analysis of 6 kb element-encoded rRNAs. Total *P. falciparum* RNA (3 μg/lane) was electrophoresed on a denaturing 12% acrylamide gel, electroblotted to nylon membrane and probed with a ^{32}P-labeled *in vitro* antisense transcript complementary to rRNA fragment ssu6A (A) and an endlabeled oligonucleotide complementary to lsu6F (B). Transcript sizes were estimated from an RNA ladder and are indicated in nt.

Analysis of rDNA fragments encoded by the 6 kb element

The cumulative lengths of lsu6 and ssu6 minimal rDNA fragments, 822 and 429 nt respectively, are substantially smaller than other rRNAs; the smallest previously reported are the mitochondrial rRNAs of kinetoplastid protozoa, averaging 1150 and 610 nucleotides, whereas those of *E. coli* are 2904 and 1542 nt (22). Despite the small size of the 6 kb rDNA fragments, however, they all encompass portions of the conserved core sequences critical for ribosome function (Figs. 4 and 5). Comparison of the potential secondary structures of the rRNAs shows that all but one of the lsu core regions have corresponding lsu6 sequences (Fig. 6A). The representation of core sequences in ssu6 is less complete since no sequences corresponding to the ssu 5′ end and the 790 loop core sequences have been identified (Fig. 6B). However, a total of over 1 kb of sequence in several small regions remains unassigned (Fig. 2) and may encode additional small rRNA fragments.

In contrast to the 6 kb element, the 35 kb circular DNA of *P. falciparum* encodes contiguous rRNAs with sizes and potential secondary structures quite similar to those predicted for *E. coli* (10, Wilson *et al.*, in preparation). Comparisons of the core region sequences of ssu6 and lsu6 sequences with the corresponding sequences from the 35 kb DNA (ssu35 and lsu35) and *E. coli* (ssuEc and lsuEc) indicate that most of the fragments are on average 50–60% conserved relative to both *E. coli* and 35 kb rRNAs (Table 1) while conservation outside the core

```
A.
  con          g·AuuuGcUa ··uuC·GUgC CAGCagCaGC GGUaAUAC·· ·gGAagaugC aAGcgUUAuu C····uUauu aGGCGU··AA AGCGuucg
  ssuEc(500)   GCACCGGCUA A-CUCCGUGC CAGCAGCCGC GGUAAUAC-- -GGAGGGUGC AAGCGUUAAU CGGAAUUACU GGGCGU--AA AGCGCACG
  ssu35(501)   GUAUUUGCUA WUUUCUGUGC CAGCAGCAGC GGUAAUAC-- -AGAAAAUGC AAGCGUUAUU C-AUUUUAUU AGGCGU--AA AGCGUUUU
  ssu6A(2004c)   AUUUGAUC CAUACAGUCC CAGCGACAGC GGUUAUACUU UGGAAGAGUC GAGUAUUAUC CAUCCAUGUC AGGCGUUAAA AGCGUUCG
                                              ↑
                                              S
B.
  con          cGCCUGgggga gUAcguCCgc A·AG·uu·AA a·uCAAa·gG AAuUGacgGg aGC··acACa AGcGGUGGAa CAugUgGuUU aAUUcGAUgc aAcaCgauaa
  ssuEc(880)   CGCCUGGGGA GUACGGCCGC A-AGGUUAAA ACUCAAA-UG AAUUGACGGG GGCCCGCACA AGCGGUGGAG CAUGUGGUUU AAUUCGAUGC AACGCGAAGA
  ssu35(876)   CGCCUGAGUA GUAUAUUCGC A-AGAAUGAA AUUCAAA-GG AAUUGACGGG AGCUUAUACA AGUGGUGGAA CAUGUGGCUU AAUUCGAUGC AACACGAUAA
  ssu6B(500c)  UGCCUGGAGG UUACGUCCAU ACAGUUAUAA G--CAAGUGG AA-UGUUAGA AGCAAACACU AGCGGUGGAA CACAUUGUUU CAUUUGAUAG UAAACACUAU
                        ↑                          ↑
                        T                          S
  con          ACCUUACCaa ··uuU·acA
  ssuEc        ACCUUACCUG GUCUUGACA
  ssu35        ACCUUACCAA AAUUUAACA
  ssu6B        ACCUUACCAA UCUAUUUGA
                          1
C.
  con          uc·agggcua c·cacgUG·c Uaca··ggc· ·uUAcaa·a· u·u·       uu·auu·auA Gu·c··GAuu guaaacUgaA acUcg·cU·c augAAGu·Gg
  ssuEc(1216)  ACCAGGGCUA CACACGUG-C UACAAUGGCG CAUACAAAGA GAAG(31 nt)GUGCGUCGUA GUCCG-GAUU GGAGUCUGCA ACUCGACUCC AUGAAGUCGG
  ssu35(1169)  UUUUGGGCUG CUCACGUGU- UACAUAAAAU AUUACAAUAU UUUA(26 nt)UAUAUUUAUA GUUCA-GAUU AUAAAUUGAA ACUCAUUUAU AUAAAGAUGG
  ssu6C(102c)  UCAAAUUGGA UGGUGUUGGC UG----GGCA UUUAAUCCAC UCUU------ UUAAUGAAAA GGAUUUGACG GUCAACUCAA UAUUG
  ssu6D(5444c)                                                                                          GCUGC UGGAAGUAGA
  con          AAUCgcUAgU AAuCGu·aau cAGaau···a ·CGgUGAAUa ·guuc····G CcUUGUACAC ACCGCcCGUC ACaccau·ga agug            ···uu
  ssuEc        AAUCGCUAGU AAUCGUGGAU CAGAAUGCCA -CGGUGAAUA CGUUCCCGGG CCUUGUACAC ACCGCCCGUC ACACCAUGGG AGUG    (38 nt)  CGCUU
  ssu35        AAUCACUAGU AAUCGCUAAU CAGAAUUAUA GCGGUGAAUA AGUUCUUAAG CUUUGUACAC ACCGCCCGUC ACAUCUAAAA AGUA    (26 nt)    AUA
  ssu6D        AAUCGUUAUU AAGCGUCAGG AAGUCC-UGG ACGUUGAAU
  ssu6E(1649)                                                    UUG CCUUGUACAC ACCGCUCGUC ACGCA
  ssu6F(5527c)                                                                   2 3   4   ↑↑   AUCGA CUCGU----- -----GAGGU
                                                                                       ↑   KP
                                                                                       K
  con          A·aA·au··· ··u··Auaau ugggguGAAG UCGUAACAaG GUAgccGua· ggGAAc·UGc gGcUGgau·a
  ssuEc        ACCACUUUGU GAUUCAUGAC UGGGGUGAAG UCGUAACAAG GUAACCGUAG GGGAACCUGC GGUUGGAUCA
  ssu35        AUAAUAUAUA AUUAUAUAAU UUAGAUGAAG UCGUAACAAG GUAGCCGUAC UGGAAGGUGC GGCUGGAUAA
  ssu6F        AAAAGAA--- -----ACAGU CGGUGCGAAG UCGUAACAUG GUAGUUGACA GUGAACUUGU AGCUG
                                            ↑4     3 2
                                            H
```

Figure 4. Comparison of small subunit rDNA sequences from the 6 kb element and 35 kb circular DNA of *P. falciparum* and *E. coli*. Sequences from the 35 kb circular DNA (ssu35) and *E. coli* (ssuEc) corresponding to the 6 kb element ssu rDNA fragments (ssu6A-F) are aligned with it and each other for maximum homology with adjustments to reflect potential conserved secondary structures. Regions identified as core sequences (47) are underlined in the *E. coli* sequence and the starting nucleotide in each set of comparisons is identified with its numerical position in the *E. coli* sequence or the *P. falciparum* database entries (ssu35, X57167; ssu6, M76611); c indicates sequence from the complementary strand. The consensus line (con) uses capital letters for sequences conserved in all three DNAs, lower case letters for those conserved in only two DNAs, and dots to indicate no conservation. Dashes in the other lines indicate gaps included to maximize homology. Positions involved in postulated long range interactions (25) are marked with numbers below the ssu6 line to indicate the appropriate matches; site 1 has only one of the pair indicated, as the other is probably in the single-stranded region at the end of ssu6D (not shown). Sites implicated in tRNA binding are shaded in the *E. coli* sequence and in the *P. falciparum* sequences, if conserved. Sites implicated in antibiotic interactions are indicated with lettered arrows. A, anisomycin; C, chloramphenicol; E, erythromycin; H, hygromycin; K, kanamycin; L, lincomycin; P, paromomycin; S, streptomycin; T, tetracycline; Th, thiostrepton.

regions is noticeably lower (Figs. 4, 5). The two principal exceptions are ssu6C and ssu6E. The former is only 11% conserved between all three rRNAs. The core sequence encompassed by ssu6C is only 28 nt and consists largely of one arm of a helix. In this case, the potential secondary structure is maintained by compensatory base changes in the 3′ end of the ssuB sequence which forms the other arm of the helix. The second exception, ssu6E, is 75% conserved between all three sequences. This is a short sequence, entirely within core sequence, and encompasses a 13 nt sequence which is almost universally conserved among small subunit sequences. Unexpectedly, the ssu6E sequence varies from ssuEc and ssu35 at one site within the 13 nt region, replacing the almost universally conserved C with a U (Fig. 4). This alteration is also seen in some fungal ssu rRNAs (23) but its significance is unknown.

Both *P. falciparum* extrachromosomal DNAs are AT-rich; the sequenced portions of the 35 kb DNA average 78% and the 6 kb element averages 68% AT. While the core rDNA regions analyzed in Figs. 4 and 5 are not as AT-rich (63% and 58% for the 35 kb and 6 kb DNAs, respectively), *E. coli* averages less than 48% AT content in the same regions. By far the majority of positions which are conserved between the 6 kb element and 35 kb circular DNA but not the *E. coli* rDNAs are AT matches (Table 2). This AT bias is distinctly different from the other comparisons and suggests that a significant proportion of the observed similarity between rDNA sequences from the two *P. falciparum* DNAs results from bias in nucleotide composition.

The maintenance of the helix formed between ssu6B and ssu6C, described above, is one of several examples of conservation of

Table 2. Nucleotide composition of conserved rRNA sequences[1].

		numbers and % conservation of sites			
rRNA	nt	Ec/35/6	Ec/35	Ec/6	6/35
ssu					
	A+T	79 = 51%	36 = 47%	8 = 24%	23 = 88%
	G+C	77 = 49%	41 = 53%	26 = 76%	3 = 12%
lsu					
	A+T	177 = 54%	56 = 56%	17 = 30%	42 = 88%
	G+C	149 = 46%	44 = 44%	39 = 70%	6 = 12%

Ec, *E. coli* rRNAs; 35, 35 kb-encoded rRNAs; 6, 6 kb-encoded rRNAs.
[1]analysis includes only core sequences presented in Figs. 4 and 5.

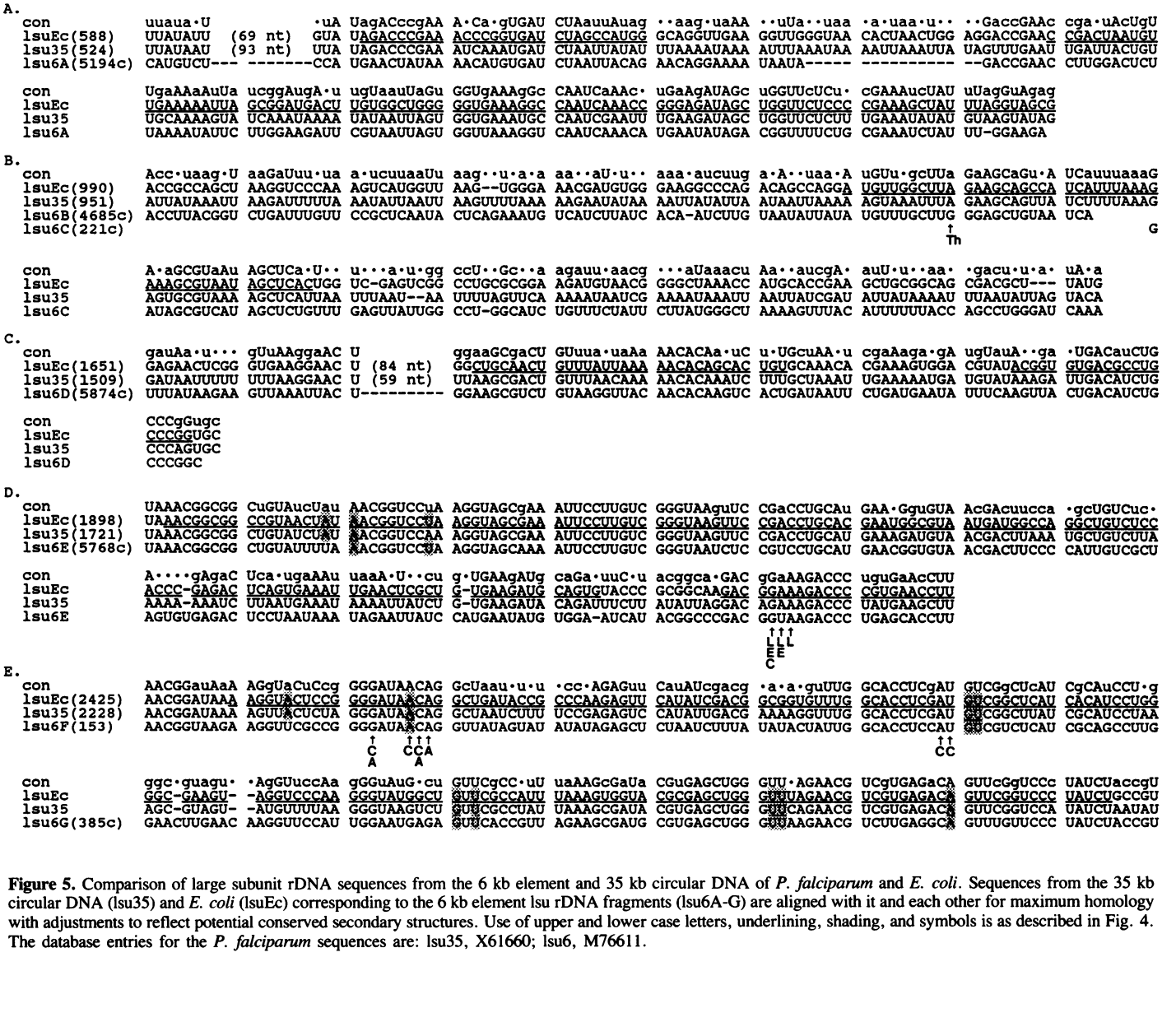

Figure 5. Comparison of large subunit rDNA sequences from the 6 kb element and 35 kb circular DNA of *P. falciparum* and *E. coli*. Sequences from the 35 kb circular DNA (lsu35) and *E. coli* (lsuEc) corresponding to the 6 kb element lsu rDNA fragments (lsu6A-G) are aligned with it and each other for maximum homology with adjustments to reflect potential conserved secondary structures. Use of upper and lower case letters, underlining, shading, and symbols is as described in Fig. 4. The database entries for the *P. falciparum* sequences are: lsu35, X61660; lsu6, M76611.

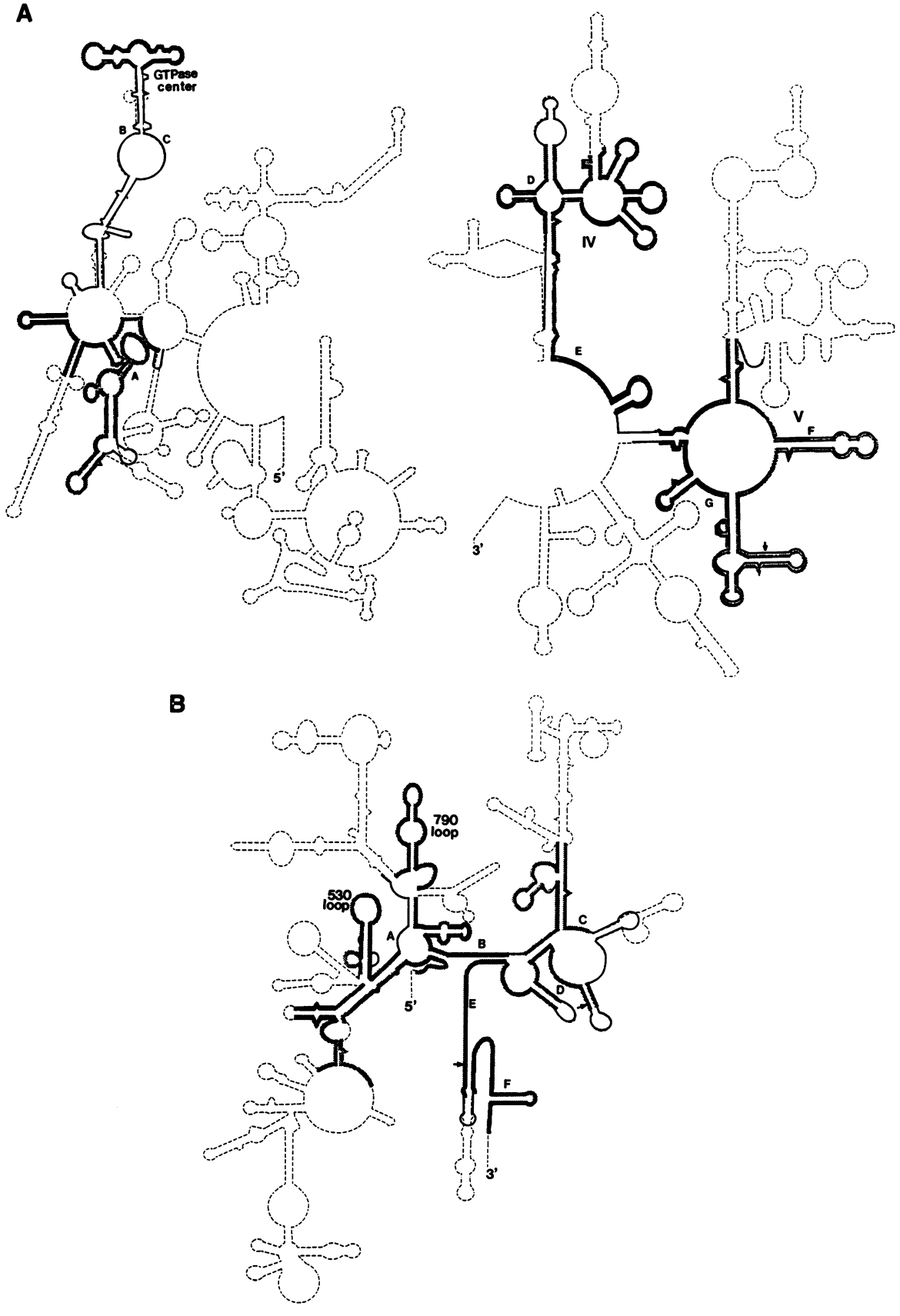

Figure 6. Potential secondary structure comparison of the 6 kb element and 35 kb circular DNA rRNAs. The predicted secondary structures of the large (A) and small (B) rRNAs encoded by the 35 kb molecule are shown by dotted lines. These secondary structures are very similar to those of *E. coli* rRNAs (10, Wilson *et al.*, in preparation) and the core regions (47) are indicated with thick gray lines. The potential secondary structures for the *P. falciparum* 6 kb element rRNA fragments are shown as thin black lines, with each fragment identified by letter. The ssuC-ssuD, ssuE-ssuF, and lsuF-lsuG junction sites are indicated with small arrows.

secondary structure in the absence of primary sequence conservation. Another example is provided by lsu6B and C (Fig. 5), the two fragments least well conserved between *P. falciparum* and *P. yoelii* (Table I). The two fragments form a long helix with two short intramolecular helices adjoining each other (Fig. 6A). The low level of conservation reflects the fact that only the short helices, which form the GTPase center, are within the core region. The arms of the longer intermolecular helix presumably have fewer constraints on evolutionary change as long as the general secondary structure is maintained. The helix and loop formed by the middle portion of lsu6F also maintain secondary structure in the absence of sequence conservation. The potential secondary structure is retained although only 15 of 42 nt are conserved between lsuEc and lsu6F in this region; in contrast, 28 of 42 nt from the corresponding lsu35 sequence match lsuEc. On the other hand, the secondary structure of the ssu 530 loop region forms a generally conserved functional pseudoknot (24), which is absent from the 6 kb element sequence (Fig. 6B).

In addition to the core regions, analysis of numerous ssu rRNAs has outlined a number of long distance interactions (25). Some of these, while maintained in ssu35, correspond to sequences which are not within the ssu fragments identified on the 6 kb element. However, three interactions near the 3′ end of ssu rRNA are conserved in all three genomes (Fig. 4) and also in *P. yoelii* (data not shown); all are in highly conserved regions near the 3′ end of ssu rRNA. One additional case of long range interactions is especially interesting as it appears to involve compensatory base changes. The interaction is between *E. coli* nt A(994) and U(1380); these positions are conserved in ssu35. The corresponding sequence for nt 994 is in ssu6B and is a U. Position 1380 is in a single-stranded region just 3′ of ssu6D and so is not within the minimum sequences in Fig. 4. However, if the ssu6D rRNA extends four nt further, the position corresponding to *E. coli* 1380 would be an A, allowing interaction with the U in ssu6B. Preliminary experiments suggest the ssu6D transcript is larger than its predicted minimum size (data not shown), suggesting that the long range interaction is maintained by compensatory transversions. Numerous tRNA binding sites have also been identified in *E. coli* rRNAs (26); the majority of those which are within the regions defined by the 6 kb rDNA fragments are conserved between *E. coli* and the *P. falciparum* 6 kb element (Figs. 4 and 5).

A number of sites which are implicated in drug resistance have been identified in both the large and small rRNAs (27). The majority of such sites which fall within the regions included in the 6 kb rDNA fragments are conserved relative to other rRNAs. These include sites for streptomycin, tetracycline, paromomycin, kanamycin, hygromycin, and aminoglycosides in ssu6 (Fig. 4). The lsu6 fragments contain sites for thiostrepton, chloramphenicol, lincomycin, and erythromycin interactions. Two sites vary from the expected sensitive sequence. One, in ssu6A (Fig. 4A), substitutes a G for an A residue (*E. coli* nt 523) correlated with streptomycin sensitivity; it is not known if such a change would result in drug resistance. A site (Fig. 5D) in the lsu peptidyltransferase domain (domain V) also varies. Sensitivity to erythromycin has been correlated with an A at the position corresponding to *E. coli* nt 2059 (27), within the single stranded region of the peptidyltransferase domain. Both the ssuEc and ssu35 sequences have an A at that site but the corresponding nucleotide in ssu6E is a U; this change is associated with erythromycin resistance in *E. coli* (27). Tetracycline, erythromycin, and chloramphenicol are known to be active against *P. falciparum in vitro*, whereas kanamycin and streptomycin are essentially inactive (28,29). The ability of these antibiotics to enter infected erythrocytes may affect their ability to block *P. falciparum* ribosome function.

DISCUSSION

Plasmodium species contain two extrachromosomal genomes, the 6 kb element and the 35 kb circular DNA. The only sequences found to be in common between them are rDNA sequences. While the 35 kb DNA encodes complete contiguous rRNAs, the 6 kb element rDNA sequences are fragmented and abbreviated. Comparison of the *P. falciparum* 6 kb element rDNA regions with corresponding sequences from *E. coli* and the 35 kb DNA shows that 50−60% of core sequences are conserved. Both the lsu and ssu rRNA core sequences from the 35 kb DNA are more conserved relative to *E. coli* than are corresponding 6 kb sequences. The core sequences from both *P. falciparum* DNAs have markedly higher AT contents than corresponding *E. coli* sequences. If, as seems likely, the base composition of the *P. falciparum* sequences has inflated the percent match between them, both the 35 kb and 6 kb rRNA sequences are more conserved relative to *E. coli* than to each other. The evolutionary implications of pairwise comparisons such as these must be interpreted carefully, as there are so few sequences being analyzed. Nevertheless, the differences between the sequences are consistent with a relatively ancient divergence between the 6 kb and 35 kb rRNAs. The apparent lack of other shared sequences between these two DNA molecules makes it highly unlikely that they have a recent common progenitor or that the 6 kb element is derived from the 35 kb circular DNA by recombination. Thus, while the 6 kb element appears to be mitochondrial, the role of the 35 kb circular DNA remains enigmatic.

The highly fragmented rRNA genes of the 6 kb element are among the most unusual yet described, with the fragment order 'scrambled' along the DNA molecule. The scrambled order and the sequences of the rDNA fragments are conserved between *P. falciparum* (this paper) and *P. gallinaceum* (4). Four rRNA fragments have been identified in *P. yoelii* (1,21) and the sequence conservation between 6 kb elements of these species suggest the others discussed here are also present. Many rRNAs are fragmented to some degree: a 5.8S rRNA often provides the 5′-most lsu sequences and plant lsu rRNAs frequently have a 4.5S rRNA containing the 3′-most sequences (30). Several lower eukaryotes have even more fragmented rRNAs. The cytoplasmic lsu rRNAs of *Crithidia fasciculata* (31) and *Euglena gracilis* (32) have seven and 16 separate components, respectively, while *Tetrahymena pyriformis* has two mitochondrial ssu modules (33). The *Chlamydomonas reinhardtii* mitochondrial lsu and ssu rRNAs appear most similar to the *Plasmodium* case, as the rDNA fragments are interspersed with each other and with some protein coding genes (34).

In all these cases, the fragmented rRNAs have the potential, by intermolecular interactions, to form secondary structures much like those proposed for continuous rRNAs. The potential secondary structures formed by the *P. falciparum* 6 kb element rRNA fragments are generally similar to those proposed for *P. gallinaceum* (4) and *P. yoelii* (1,21), although the details, especially of the potential helix between lsu6B and lsu6C, differ. The 6 kb fragments differ from other rRNAs in that substantial

portions of rRNA structure are apparently missing. Intriguingly, the majority of the ssu6 and lsu6 sequences correspond to regions of rRNA which are highly conserved core sequences, including sites which are implicated in ribosome function. Sequences corresponding to most of the variable regions, which in other organisms often still form similar secondary structures despite lack of sequence conservation, have not been found. The positioning of breaks in the discontinuous ssu6 and lsu6 sequences also differs from other fragmented rRNAs. The breaks between rDNA sequences in *C. fasciculata*, *E. gracilis* and *C. reinhardtii* are generally located in variable regions (31,32,34,35). In contrast, some of the discontinuities between 6 kb element-encoded rRNA pieces occur in conserved regions, including the single stranded region between ssu6D and E, the GTPase center defined by lsu6B and C, and the helical region near the 3′ end of the lsu rRNA which includes the 3′ end of lsu6F and the 5′ end of lsu6G.

Some core sequences and other regions which have been implicated in important ribosomal functions have not yet been identified in the 6 kb element. Such missing regions include the lsu a-sarcin loop which functions in binding of ternary complexes (36) and ssu regions implicated in ribosome assembly or protein synthesis (37–41). It seems likely that at least some of these functionally important regions exist as other rDNA modules in the 6 kb element but have not yet been identified, perhaps because very small size or lack of sequence conservation makes them difficult to detect. Alternatively, they might be encoded by the nuclear or 35 kb genomes; import of small RNAs into the mitochondrion has been documented (42,43). Some of the missing regions may be functionally unnecessary. For example, in the *E. coli* ssu rRNA, position 1054 and the region between nt 1199 and 1204 are important for recognizing UGA as a termination codon (30,37,39). The 35 kb DNA ssu sequence has conserved these sites but the corresponding region is absent from the 6 kb element. However, none of the 6 kb element protein coding genes employ UGA as a stop codon, nor do they use it as a tryptophan codon, as many mitochondrial genomes do. In effect, the ability to distinguish proper use for UGA codons may not be required for the 6 kb element, making that region of rRNA sequence perhaps less critical for survival.

The close correspondence between the 6 kb element rRNA pieces and rRNA conserved core regions suggests that these fragments may provide the minimalized rRNA portions of functional ribosomes. Transcripts of the rDNA fragments are abundant, as would be expected for rRNAs; similar abundant small transcripts have been reported from the *P. gallinaceum* (2) and *P. yoelii* (21) 6 kb elements. Functionally important tRNA and drug interaction sites are generally conserved. Furthermore, the degree of conservation of the 6 kb element rDNA sequences between *Plasmodium* species exceeds the degree of nucleotide sequence conservation for the protein coding genes and approximates that of their amino acid sequence conservation, implying that there is evolutionary pressure to preserve these sequences. The element encodes cytochrome genes whose products must be expressed since cytochrome oxidase activity has been detected in malaria parasites (44–46). Ribosomal particles containing 6 kb element-encoded rRNAs would be obvious candidates for translating the cytochrome mRNAs. Complete assessment of the functionality of the rRNA fragments encoded by the *P. falciparum* 6 kb element awaits investigation of ribosomal particles.

ACKNOWLEDGEMENTS

We thank Drs. Jeffrey Joseph and Dyann Wirth for providing proposed secondary structures for the *P. gallinaceum* 6 kb element-encoded rRNA fragments, Dr. Akhil Vaidya for the *P. yoelii* 6 kb element clone, Peter Moore for initial cloning of the *P. falciparum* 6 kb element, John Sumpter for assistance in preparing Fig. 6, Karen Kinch for secretarial help, and Dr. Marilyn Parsons for critical comments on the manuscript. This work was supported by NIH grant AI25513 and NATO grant RG.0132/88 to J.E.F. and a UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases (TDR) grant to R.J.M. Wilson.

REFERENCES

1. Vaidya,A.B., Akella,R. and Suplick,K. (1989) Mol. Biochem. Parasitol., **35**, 97–108. Corrigendum (1990) Mol. Biochem. Parasitol. **39**, 295–296.
2. Aldritt,S.M., Joseph,J.T. and Wirth,D.F. (1989) Mol. Cell. Biol., **9**, 3614–3620.
3. Joseph,J.T., Aldritt,S.M., Unnasch,T., Puijalon,O. and Wirth,D.F. (1989) Mol. Cell. Biol., **9**, 3621–3629.
4. Joseph,J.T. (1990) The mitochondrial microgenome of malaria parasites. Ph.D. thesis, Harvard Medical School.
5. Gray,M.W. (1989) Annu. Rev. Cell Biol., **5**, 25–50.
6. Gardner,M.J., Bates,P.A., Ling,I.T., Moore,D.J., McCready,S., Gunasekera,M.B.R., Wilson,R.J.M. and Williamson,D.H. (1988) Mol. Biochem. Parasitol., **31**, 11–17.
7. Gardner,M.J., Williamson,D.H. and Wilson,R.J.M. (1991) Mol. Biochem. Parasitol., **44**, 115–124.
8. Feagin,J.E., Gardner,M.J., Williamson,D.H. and Wilson,R.J.M. (1991) J. Protozool., **38**, 243–245.
9. Wilson,R.J.M., Gardner,M.J., Williamson,D.H. and Feagin,J.E. (1991) Parasitol. Today, **7**, 134–136.
10. 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.
11. Fox,T.D. (1983) Nature, **301**, 371–372.
12. Thorsness,P.E. and Fox,T.D. (1990) Nature, **346**, 376–379.
13. Hempelmann,E., Ling,I. and Wilson,R.J.M. (1981) Trans. R. Soc. Trop. Med. Hyg., **775**, 855–858.
14. Trager,W. and Jensen,J.B. (1978) Nature, **273**, 621–622.
15. Tungpradabkul,S. and Panyim,S. (1985) In Panyim,S., Wilairat,P. and Yuthavong,Y. (eds.), Application of genetic engineering to research on tropical disease pathogens with special reference to Plasmodia. World Health Organization (TDR), Geneva, pp. 51–56.
16. Milhausen,M., Nelson,R.G., Parsons,M., Newport,G., Stuart,K. and Agabian,N. (1983) Mol. Biochem. Parasitol., **9**, 241–254.
17. Chomczynski,P. and Sacchi,N. (1987) Anal. Biochem., **162**, 156–159.
18. Feinberg,A.P. and Vogelstein,B. (1983) Anal. Biochem., **132**, 6–13.
19. Devereux,J., Haeberli,P. and Smithies,O. (1984) Nucl. Acids Res., **12**, 387–395.
20. Williamson,D.H., Wilson,R.J.M., Bates,P.A., McCready,S., Perler,F. and Qiang,B. (1985) Mol. Biochem. Parasitol., **14**, 199–209.
21. Suplick,K., Morrisey,J. and Vaidya,A.B. (1990) Mol. Cell. Biol., **10**, 6381–6388.
22. Benne,R. and Sloof,P. (1987) BioSystems, **21**, 51–68.
23. Dams,E., Hendricks,L., Van de Peer,Y., Neefs,J.M., Smits,G., Vandembempt,I. and De Wachter,R. (1988) Nucl. Acids Res., **16(Suppl)**, r87–r173.
24. Powers,T. and Noller,H.F. (1991) EMBO J., **10**, 2203–2214.
25. Haselman,T., Camp,D.G. and Fox,G.E. (1989) Nucleic Acids Res., **17**, 2215–2222.
26. Noller,H.F., Moazed,D., Stern,S., Powers,T., Allen,P.N., Robertson,J.M., Weiser,B. and Triman,K. (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. 73–92.
27. Cundliffe,E. (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. 479–490.

28. Geary,T.G. and Jensen,J.B. (1983) Am. J. Trop. Med. Hyg., **32**, 221–225.
29. Divo,A.A., Geary,T.G. and Jensen,J.B. (1985) Antimicrob. Agents. Chemother., **27**, 21–27.
30. Hanfleur,A., Kleuvers,B. and Göringer,H.U. (1990) Nucl. Acids Res., **18**, 5625–5632.
31. Spencer,D.F., Collings,J.C., Schnare,M.N. and Gray,M.W. (1987) EMBO J., **6**, 1063–1071.
32. Schnare,M.N. and Gray,M.W. (1990) J. Mol. Biol., **215**, 73–83.
33. Schnare,M.N., Heinonen,T.Y.K., Young,P.G. and Gray,M.W. (1986) J. Biol. Chem., **261**, 5187–5193.
34. Boer,P. and Gray,M.W. (1988) Cell, **55**, 399–411.
35. Schnare,M.N., Cook,J.R. and Gray,M.W. (1990) J. Mol. Biol., **215**, 85–91.
36. Tapprich,W.E. and Dahlberg,A.E. (1990) EMBO J., **9**, 2649–2655.
37. Dahlberg,A.E. (1989) Cell, **57**, 525–529.
38. Prescott,C.D. and Dahlberg,A.E. (1990) EMBO J., **9**, 289–294.
39. Tapprich,W.E., Göringer,H.U., de Stasio,E., Prescott,C. and Dahlberg,A.E. (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. 236–242.
40. Prescott,C.D. and Göringer,H.U. (1990) Nucleic Acids Res., **18**, 5381–5386.
41. Santer,M., Bennett-Guerrero,E., Byahatti,S., Czarnecki,S., O'Connell,D., Meyer,M., Khoury,J., Cheng,X., Schwartz,I. and McLaughlin,J. (1990) Proc. Natl. Acad. Sci. USA, **87**, 3700–3704.
42. Simpson,A.M., Suyama,Y., Dewes,H., Campbell,D.A. and Simpson,L. (1989) Nucleic Acids Res., **17**, 5427–5445.
43. Chang,D.D. and Clayton,D.A. (1989) Cell, **56**, 131–139.
44. Scheibel,L.W. and Miller,J. (1969) J. Parasitol., **55**, 825–829.
45. Scheibel,L.W. and Pflaum,W.K. (1970) J. Parasitol., **56**, 1054.
46. Fry,M. and Beesley,J.E. (1991) Parasitology, **102**, 17–26.
47. Cedergren,R., Gray,M.W., Abel,Y. and Sankoff,D. (1988) J. Mol. Evol., **28**, 98–112.