Primary and secondary structures of *Escherichia coli* MRE 600 23S ribosomal RNA. Comparison with models of secondary structure for maize chloroplast 23S rRNA and for large portions of mouse and human 16S mitochondrial rRNAs

Christiane Branlant*, Alain Krol, Mohamed Ali Machatt, Jean Pouyet and Jean-Pierre Ebel
Institut de Biologie Moléculaire et Cellulaire du CNRS, Laboratoire de Biochimie, 15 rue R.
Descartes, 67084 Strasbourg Cedex, France, and

Kaylene Edwards and Hans Kössel
Institut für Biologie III, Universität Freiburg, Schänzlestrasse 1, 7800 Freiburg, GFR

Received 2 July 1981

ABSTRACT

We determined 90% of the primary structure of *E.coli* MRE 600 23S rRNA by applying the sequencing gel technique to products of T1, S1, A and *Naja oxiana* nuclease digestion. Eight cistron heterogeneities were detected, as well as 16 differences with the published sequence of a 23S rRNA gene of an *E.coli* K12 strain. The positions of 13 post-transcriptionally modified nucleotides and of single-stranded, double-stranded and subunit surface regions of *E.coli* 23S rRNA were identified. Using these experimental results and by comparing the sequences of *E.coli* 23S rRNA, maize chloro. 23S rRNA and mouse and human mit 16S rRNAs, we built models of secondary structure for the two 23S rRNAs and for large portions of the two mit rRNAs. The structures proposed for maize chloroplast and *E.coli* 23S rRNAs are very similar, consisting of 7 domains closed by long-range base-pairings. In the mitochondrial 16S rRNAs, 3 of these domains are strongly reduced in size and have a very different primary structure compared to those of the 23S rRNAs. These domains were previously found to constitute a compact area in the *E.coli* 50S subunits. The conserved domains do not belong to this area and contain almost all the modified nucleotides. The most highly conserved domain, 2042-2625, is probably part of the ribosomal A site. Finally, our study strongly suggests that in cytoplasmic ribosomes the 3'-end of 5.8S rRNA is base-paired with the 5'-end of 26S or 28S rRNA. This confirms the idea that 5.8S RNA is the counterpart of the 5'-terminal region of prokaryotic 23S rRNA.

INTRODUCTION

We have previously sequenced fragments of *E.coli* MRE 600 23S rRNA encompassing about 75% of the molecule (1). The relative order of these fragments can now be unambiguously determined by comparison with the sequence established for a 23S RNA gene of an *E.coli* K12 strain (2). As differences were observed between some of the sequenced fragments and their counterparts in the rDNA (1, 3,4), we decided to complete our sequence analysis of *E.coli* MRE 600 23S RNA. Since the experiments performed for this purpose also gave us information on the secondary structure, we began a systematic study of this secondary structure.

The secondary structures of 5S (5) and 16S rRNAs (6,7,8) have already been studied and were found to be highly conserved throughout evolution, base-pair-

ings having been maintained by compensating mutations. For instance, a secondary structure similar to that established for E.coli 16S RNA can be formulated for maize chloroplast 16S RNA (8). We postulated a similar secondary structure conservation for 23S RNA. Since genes coding for maize chloroplast 23S RNA (9) and for mouse and human mitochondrial 16S large rRNAs (10,11) have been sequenced, we compared the possible base-pairings in these three RNAs with those in E.coli 23S RNA. Using these results and those obtained in the course of the sequence study, we could propose a model of secondary structure for E.coli and maize 23S RNAs and for large portions of the two mit 16S RNAs. As fragments of large subunit rRNA from several other organisms have been sequenced (12-19), we also looked for possible secondary structure homologies between these fragments and E.coli and maize chloroplast 23S RNAs.

In this paper we describe both our experimental study of E.coli MRE 600 23S RNA primary and secondary structures and our comparison of sequence complementarity in large rRNAs. The results obtained are discussed, taking into account published data on the structure of 50S subunits and on putative functional sites of 23S RNA.

MATERIALS AND METHODS

Preparation of E.coli MRE 600 23S RNA and 50S subunits was as previously described (20).

1. Partial digestion of 50S subunits was i) with Naja oxiana RNase (21), in 0.3 M KCl, 10 mM MgCl₂, 10 mM Tris-HCl pH 7.5 buffer, with 8 U of RNase per 100 µg of subunits, for 4 hours at 0°C ; ii) with T1 RNase, in 10 mM MgCl₂, 10 mM Tris-HCl pH 7.5 buffer, in the presence or absence of 2 M urea, for half an hour at 0°C, at enzyme/substrate ratios between 1/100 and 1/1000 (w/w) ; iii) with S1 nuclease in 3 M NaCl, 0.001 M ZnCl₂, 0.05 M Na acetate buffer pH 5.7, overnight at 4°C, with 250 U of S1 nuclease per 50 µg of subunit. In each case, the digestion fragments were extracted with phenol and fractionated on a 6% polyacrylamide gel.

2. Partial digestion of 23S RNA was in 10 mM Tris-HCl pH 7.5 buffer in the presence or absence of 10 mM MgCl₂, for half an hour at 0°C, with T1 or pancreatic RNase. The enzyme/substrate ratios were between 1/3000 and 1/1000 (w/w) for T1 RNase and 1/100 000 and 1/10 000 (w/w) for pancreatic RNase. The digestions were stopped by phenolic extraction and the fragments were fractionated on 12% polyacrylamide gels.

3. Sequence analysis of the resulting fragments : 5'-ends were labeled with γ-³²P ATP and T4 polynucleotide kinase and 3'-ends with 5'-³²P pCp and T4 RNA

ligase. When necessary, the terminal phosphate was eliminated prior to labeling by the action of alkaline phosphatase. 5'-end labeled fragments were analyzed by the enzymatic method for RNA sequencing (1), 3'-end labeled fragments by both the enzymatic and chemical methods (22).

4. Computer analysis of complementary sequences was done on a minicomputer using a program derived from that of Pipas and Mc Mahon (23).

RESULTS

1. Experimental study of E.coli MRE 600 23S RNA. The 50S subunits and isolated 23S RNA were digested under a variety of conditions with various RNases as described in Methods. T1 and S1 nucleases were used in conditions such that they cleaved single-stranded regions, whereas *Naja oxiana* nuclease was used in conditions where it cleaved only base-paired RNA regions (24,25). The fragments resulting from these digestions were end-labeled and sequenced on gels using either the enzymatic method or the chemical method for partial digestions.

1.1 The primary structure of E.coli MRE 600 23S RNA. The sequenced fragments cover 90% of the molecule (Fig. 1). We failed to isolate pure fragments from the remaining 10% of the molecule. The sequence obtained contains 8 microheterogeneities which are probably due to sequence differences in the 7 genes coding for 23S RNA (27). These heterogeneities are indicated in Figures 2 to 6. In addition, this sequence differs at 16 positions from that of the rDNA of *E.coli* K12. Seven of these differences were already described in previous papers dealing with the sequence of the protein L1 binding region (3) or with the structure of the 3'-terminal region of 23S RNA (4). The additional differences observed are located in the central region of the molecule : U1542 is deleted, G1723 and C1730 are replaced by A and U respectively and the segment 1171-1178 (G-C-U-U-A-U-G-Cp) is replaced by A-U-C-A-G----Up (the nucleotides 1176 and 1177 having not been identified). Interestingly, the segment 1171-1178 probably contains the very sensitive phosphodiester bond of 23S RNA whose cleavage gives rise to the 13S and 18S fragments (26). The 16 differences observed could result either from cistron heterogeneities or from strain differences.

We have already localized two post-transcriptionally modified nucleotides of 23S RNA : m⁶A and m⁷G, at positions 2030 and 2069 respectively (3). Since in the present study nucleotides 2251 and 2498 were cleaved neither with RNase nor in boiling water, we concluded that they are 2'-O methylated. Furthermore, nucleotides 745, 1618, 1915 and 2449 which were cleaved in boiling water were cut by none of the RNases. Using these results and the sequence of the 23S RNA T1 digestion products containing modified nucleotides (28,29), we were able to loca-

Figure 1 : Schematic representation of the regions of *E. coli* MRE 600 which have been sequenced.

lize 11 additional modified nucleotides. Ten of them were previously observed in 23S RNA by Fellner and Sanger (28) 3Ψ, 2T, Gm, Cm, m⁶A, m¹G and an unidentified methylated U. They are located at positions 746, 1911, 1917, 747, 1939, 2251, 2498, 1618, 745 and 1915 respectively. The eleventh is an unidentified modified U which we previously detected by fingerprint analysis and which is located at position 2449.

1.2 Identification of single-stranded and double-stranded regions in *E. coli* MRE 600 23S RNA. The extremities of the 23S RNA digestion products were identified on sequencing gels, thus indicating nucleotides in single-stranded or

GENERAL LEGEND TO FIGURES

Models of *E. coli* 23S RNA. The sequence given is that of the rDNA (2). The heterogeneities and sequence differences observed in *E. coli* MRE 600 23S RNA are represented by circled nucleotides, the post-transcriptional modifications are indicated on the sequence itself, * corresponds to the presence of a non identified post-transcriptional modification. The positions of nuclease digestion of *E. coli* 23S RNA within 50S subunits are indicated by ◀○ T1 (in the presence of urea), ◀● T1 (in the absence of urea), ◀■ S1 and ◀□ *Naja oxiana*, the position of nuclease digestion of isolated 23S RNA are indicated by ◀◊ T1 and ◀◆ panc.. K indicates the site of kethoxal modification, (K) corresponds to sites which are protected in the presence of the 30S subunits (43). When hairpin structures existed in isolated fragments and were resistant to the denaturing conditions of sequence analysis, we have indicated the extremities of the fragment in which they were observed (↗ ↖), the relative sensitivity of their guanines to T1 RNase digestion as deduced from the sequencing gel (■, ●, —) and the nucleotides corresponding to band compression on the sequencing gel (↗↖). The proposed base-pairings which are strongly supported by the experimental results and (or) by the results of sequence comparison, are indicated by a full line across the base-pairs. The others are considered as putative. Alternative base-pairings are indicated by sequences boxed by dotted lines and joined together by an arrow.

Models of maize chloroplast 23S RNA. The sequence is that established by Edwards and Kössel (9). The base-pairs which are identical in maize and *E. coli* RNAs are represented by thick bars, those represented by thin bars are either conserved by base-compensation or are additional. The nucleotides in single-strands which are conserved from *E. coli* to maize chloroplast have been boxed by full lines, the semi-conserved ones by dashed lines.

Models of mit 16S RNAs. The sequence given is that of mouse mit RNA (10). The mutations in human mit RNA (11) compared to mouse RNA have been indicated in circles, ▼ represents a deletion. Sequence conservation from *E. coli* to mit RNA is represented as for chloroplast RNA.

DOMAIN I

Figure 2 : a) The region 1-530 of *E. coli* 23S RNA. Fragments encompassing the region 10-530 were found associated with protein L24 after digestion of the L24-23S RNA complex (31). The positions of the extremities of these fragments are indicated (◄), the numbering is that previously adopted. Several of the hairpin structures were stable in the denaturing conditions of sequence analysis (sequenced fragments 1 to 4).

b) The region 1-548 of maize chloro. RNA. Compared to the corresponding region of *E. coli* RNA this region of maize RNA contains a long deletion (nucleotides 133-147 in *E. coli*) and a long insertion (nucleotides 258-279 in maize). The deletion leads to the disappearance of a hairpin structure and the insertion to the appearance of a possible additional hairpin structure.

c) Putative base-pairings of the region 1-223 of mouse mit RNA.

d) A model explaining the interaction between 5.8 S rRNA and 28S rRNA. The sequences are those of *X. laevis* rRNAs (18). The nucleotides found at identical positions in the structure of the corresponding region of *E. coli* 23S RNA are boxed.

in double-stranded regions, depending on the enzyme used for the digestion. Furthermore, the intensities and relative positions of the bands on these sequencing gels reflect the secondary structure of the sequenced fragments. For

Figure 2b

Figure 2c

Figure 2d

example, since the secondary structure was not completely destroyed under the conditions used for statistical T1 RNase digestion, the bands corresponding to cleavage at guanines in single-strands were more intense than those corresponding to cleavage at guanines in double-strands. Furthermore, some strong base-pairings were stable under the denaturing conditions used for gel electrophoresis, thus producing band compressions. This information has, however, been used with caution since structural rearrangements might occur in isolated fragments. The structures deduced were only considered valid when confirmed by sequence comparison as described below.

2. Model building. As a first step, we aligned the sequences of *E.coli* 23S RNA, maize chloroplast 23S RNA and mouse and human mit 16S RNAs for maximum sequence homology. Two regions of strong homology were observed between the four RNA species : i) the regions 229-612 of mouse and 223-601 of human mit 16S RNAs correspond to the regions 577-1195 and 594-1226 of *E.coli* and maize 23S RNAs respectively ; ii) the regions 794-1463 of mouse and 780-1440 of human mit 16S RNAs correspond to the regions 1651-2629 and 1766-2724 of *E.coli* and maize 23S RNAs. In the mit RNAs, the 3' and 5'- terminal regions, as well as the region located between the two conserved segments differ greatly in both size and sequence from the corresponding regions of the 23S RNAs. We proposed that the 23S RNA regions which are highly conserved in mit 16S RNAs and those which are poorly

Figure 3 : a) The region 531-1195 of E.coli 23S RNA. Base-pairs and single-stranded nucleotides which are conserved in both chloro. and mit RNAs are indicated by thick bars and boxed nucleotides, respectively.

b) The region 549-1226 of maize chloro. 23S RNA. The additional segment (949-970) compared to E.coli RNA can form a hairpin structure.

c) The regions 224-611 of mouse 16S mit RNA and 219-601 of human 16S mit RNA. The structure formulated for segments 439-504 and 557-594 in mouse RNA is only tentative since the corresponding structure in human RNA has different loop sizes. The mutations from mouse to man are not indicated for these segments.

Figure 3b

Figure 3c

conserved constitute independent domains of secondary structure.

When looking for the secondary structure of the 23S RNA regions which have a counterpart in mit RNAs the following strategy was adopted. The complementary sequences existing in the *E.coli* RNA region were found by computer analysis and those filling the following criteria were selected : 1) must not contain T1 or S1 cleavage sites, 2) maize RNA must possess corresponding complementary sequences, 3) in the mit RNAs, the corresponding sequences must be either complementary or deleted. Using the base-pairings selected in this way, a model of secondary structure was built for the two regions 577-1195 (Fig. 3a) and 1651-2629 (Fig. 5a) of *E.coli* 23S RNA and for the homologous regions in maize 23S RNA (Fig. 3b and 5b) and in mouse and human mit 16S RNAs (Fig. 3c and 5c). The secondary structures of these two *E.coli* RNA regions are strongly conserved in the 3 other RNA species, so that almost all the base-pairings proposed are convincingly supported by compensatory mutations.

For the three 23S RNA regions which are less conserved in mit RNAs, sequence comparison with these mit RNAs could not be made. But information on the secondary structure of these three RNA regions has been previously obtained by studies on protein-23S RNA complexes (31,32) or on the structure of *P.vulgaris* and *A.punctata* 23S RNA fragments (4).

The 5'-terminal region (1-530) of *E.coli* 23S RNA remains associated with protein L24 upon digestion of the L24-23S RNA complex (31). We have previously

demonstrated the existence of long-range RNA-RNA interactions in this area (33, 34) and can thus propose a convincing model for this region of E.coli and maize 23S RNAs (Fig. 2a and b). Several of the hairpin structures proposed were present in sequenced RNA fragments. Only the structure of the segment 446-514 is ambiguous. This segment displays few internal complementarities and among those found, two alternative structures are possible: one is consistent with base-compensation in maize RNA, the other with sequence conservation in mit RNA (Fig. 2). Indeed, the segment 446-483 of E.coli 23S RNA is strongly conserved in mit 16S RNA. In the rest of the 5'-terminal region of mit RNA the sequence has diverged, nevertheless a hairpin structure similar to that proposed for the segment 183-215 of E.coli RNA can be postulated (Fig. 2c).

The central region of E.coli 23S RNA (1196-1650) contains the binding site for protein L23 (32). The ribonucleoprotein obtained after digestion of the complex contained the segments 1340-1416 and 1588-1619, which are associated by RNA-RNA interactions (35). This observation gives support to our proposal of long-range base-pairings (Fig. 4a). However, several of the other base-pairings proposed for this RNA region are only tentative. Indeed, there are several insertions in the maize RNA region compared to the corresponding E.coli region, and several alternative base-pairings are compatible with the sequence of both RNAs.

The 3'-terminal region (2630-2904 in E.coli 23S RNA). We have previously proposed a secondary structure for the 3'-terminal 130 nucleotides of E.coli 23S RNA (4) and shown that chloroplast 4.5S RNA corresponds to this region of E.coli rRNA (4,36). The present results allow us to propose a model for the entire 3'-terminal region (Fig. 6). However, it has been possible to choose between the two alternative long-range interactions (Fig. 6) which are compatible with both 23S RNA sequences. Furthermore, one ambiguity remains concerning this part of maize RNA as its 3'-end has not been identified.

3. Comparison with the possible structure of other eukaryotic rRNA fragments.

3.1 rRNA fragments corresponding to the 5'-terminal region of 23S RNA. On the basis of sequence homology, Nazar (37) proposed that 5.8S rRNA plays the same role in eukaryotic ribosomes as the 150 5'-terminal nucleotides of 23S RNA in prokaryotic ribosomes. We indeed observed that a secondary structure very similar to that proposed for the 5'-terminal region of 23S RNA can be constructed using eukaryotic 5.8S rRNA as the counterpart to the E.coli 23S RNA sequence 1-163 and the 5'-end of the large eukaryotic rRNA for the counterpart to the E.coli 23S RNA sequence 164-270. Figure 2d shows the structure obtained for X.laevis rRNAs. A similar structure can be proposed with yeast rRNAs. In

both cases conserved nucleotide sequences are observed in the single-stranded segments. The model we propose is in good agreement with the experimental results of Pace et al. (38). However, it is not compatible with the model proposed by Kelly and Cox (39). The latter model cannot be a general one as it fits with the sequences of Neurospora crassa rRNAs only. The present results strongly suggest that the gene for eukaryotic 5.8S RNA has been separated from that of the 28S rRNA by an insertion in the ancestral gene coding for the large rRNA. The nucleotide sequence of the 5'-terminal region of Aspergillus nidulans mit 23S rRNA has also been determined (40). Like the 5'-terminal region of cytoplasmic 26 or 28S rRNAs, this mit rRNA region is similar to the region of E.coli

Figure 4 : a) The region 1196-1650 of E.coli 23S RNA. This E.coli RNA region contains the binding site for protein L23. $\leftarrow$ L23, indicates the extremities of the two segments found associated with the protein. We failed to find identical structures for the segment 1420-1471 of E.coli RNA and its counterpart in maize RNA. The base-pairings indicated with dotted lines display some similarity to those proposed for maize RNA. The structure shown in the inset was found to exist in the sequenced fragment 7. In the region of maize RNA corresponding to the segment 1521-1586 an insertion has occurred which leads to the existence of two similar sequences (1565-1604 and 1633-1668). Both could correspond to the E.coli segment 1521-1559. Different structures can be predicted for the E.coli region 1521-1586 depending on the maize RNA sequence considered as its counterpart. The structure proposed for the segment 1530-1541 was found to exist in the sequenced fragment 8.

b) The region 1227-1765 of maize chloro. 23S RNA.

Figure 5 : a) The region 1651-2629 of E.coli 23S RNA. We have indicated the extremities of the region found associated with protein L1 ($\leftarrow$ L1) (3) and those of the region found associated with 5S RNA and proteins L5, L18, L25 ($\leftarrow$ 5S) (46), the positions of the nucleotides corresponding to those conferring chloramphenicol resistance in 21S mit rRNA of yeast ($\oplus$) (15,2), the probable site of crosslinking between 23S rRNA and a puromycin derivative ($\leftarrow$ Pu) (48), the position of Um $\star$ has been deduced by Bear and Dubin (30) by comparison with the sequence containing this modified nucleotide in hamster mit RNA. Base-pairs and single-stranded nucleotides which are conserved in both chloro. and mit RNAs are indicated by thick bars and boxed nucleotides, respectively.

b) The region 1766-2724 of maize chloro. 23S RNA. The region 1926-2340 is compared to the corresponding one in Drosophila virilis 28S rRNA (14) : the hairpin structures which can be constructed with the sequence of this latter RNA are indicated by a full line across the base-pairs, the conserved nucleotides in single-strands are underlined with full lines and the semi-conserved ones with dashed lines. In the same way, the region 2440-2724 is compared to the corresponding one in yeast mit 21S rRNA (15). The phosphodiester bonds corresponding to positions where introns were observed in the rDNA of some eukaryotic rRNAs are indicated ($\leftarrow$ I), 1 Drosophila virilis 28S RNA (14), 2 Tetrahymena pigmentosa 28S RNA (12), 3 Physarum polycephalum 26S RNA (17), 4 yeast mit 21S RNA (15), 5 Chlamydomonas chloroplast 23S RNA (16).

c) The regions 793-1464 and 780-1439 of mouse and human mit 16S RNAs respectively. $\star$ these nucleotides were found to be 2'-O methylated in hamster mit RNA (30). Bear and Dubin proposed the same structure for the region 1283-1441 (30).

Figure 4a

Figure 5a

Figure 5b

Figure 5c

23S RNA starting at position 164. This suggests that an RNA similar to cytoplasmic 5.8S RNA may exist in the mitochondrial ribosomes of these primitive eukaryotes.

3.2 Eukaryotic rRNA fragments corresponding to the regions 1700-2700 of *E. coli* 23S RNA. Several rDNA fragments coding for regions of eukaryotic rRNAs corresponding to part of the *E. coli* rRNA region 1700-2700 have been sequenced. They originate from *Tetrahymena* (12), *X. laevis* (13) and *Drosophila virilis* (14) nuclear rDNA, from yeast mit rDNA (15) and from *Chlamydomonas* chloroplast rDNA (16). Secondary structures similar to those proposed for the corresponding region of *E. coli* rRNA can be formulated for all of these eukaryotic rRNA regions (Fig. 5). The *E. coli* RNA region 2150-2710 contains long single-stranded stretches which are highly conserved in all the RNAs examined (Fig. 5a,b).

Figure 6 : a) The region 2630-2904 of *E. coli* 23S RNA, b) its counterpart in maize chloro. 4.5S and 23S RNAs. In chloroplast ribosomes of flowering plants, 4.5S rRNA is the counterpart of the 130 nucleotides at the 3'-end of prokaryotic 23S RNA (4). The 3'-end of maize chloro. 23S RNA has not been identified so that the hairpin structure 2791-2805 in *E. coli* RNA may be replaced by a more complex structure in maize RNA.

NOTE : Glotz, C., Zwieb, C. and Brimacombe, R. independently used the sequence of 23S rRNA from maize chloroplast determined by Edwards and Kössel in order to establish models of secondary structures for *E. coli* and maize chloroplast 23S rRNAs. (Nucl. Acids Res., in press).

DISCUSSION

Several arguments support the validity of the models of secondary structure which we propose. a) Data from RNase digestion and from sequence comparisons agree well. Generally we did not find T1 or S1 nuclease cuts in complementary *E.coli* 23S RNA segments whose counterparts in the 3 other RNA species studied were also complementary. On the other hand, *Naja oxiana* RNase cuts often occurred in these segments. These results confirm our starting hypothesis according to which the secondary structure of prokaryotic 23S rRNA has been conserved throughout evolution. b) Conserved nucleotide sequences were found at identical positions in the secondary structure model proposed for the different rRNA species. c) In *E.coli* RNA, the post-transcriptionally modified nucleotides are in single-stranded regions and cistron heterogeneities and sequence differences with the *E.coli* K12 rDNA either compensate each other or are located in single-stranded regions. d) A secondary structure similar to that proposed for the corresponding region of 23S RNA can be constructed for all of the rRNA fragments from various eukaryotic ribosome species that have been sequenced.

Only in a limited number of regions we were unable either to show a strict conservation of the secondary structure between the four rRNA species or to propose alternative structures. In these regions the models have to be refined in the future by comparison with possible structures in other 23S RNA species.

According to the proposed models, 23S RNA contains 7 domains closed by long distance base-pairings (these domains have been denoted I to VII in the figures). Each of these domains displays sites of T1, S1 and *Naja oxiana* nuclease digestion within the 50S subunit. Therefore, each of these domains have both single-stranded and double-stranded regions at the surface of the subunit.

The 5' and 3' domains (I and VII) were previously found to be highly resistant to T1 RNase in *E.coli* 23S RNA. The stability of the secondary structure of domain I and of part of domain VII was confirmed by the sequencing gel technique. Most of the hairpin structures in domain I and some in domain VII resist the destabilizing conditions of the sequencing techniques (Fig. 2, 6). Furthermore, inspection of the gel autoradiographs also suggest the existence of interactions of a type other than the classical Watson-Crick base-pairing which occur in some of the loops (Fig. 2). Domains I and VII should be close to each other in the subunits as the 3' and 5'-ends of the molecule are base-paired. They might constitute the back side of the subunit as the 3'-end of 23S RNA has been located on this side by immune electron microscopy (41). We previously showed that these two domains, together with the central domain IV, a few other small pieces of 23S RNA and a limited number of 50S proteins consti-

tute a compact area of the 50S subunits which resists T1 RNase digestion (42). These three domains which probably have a structural role have been strongly reduced in size in mit RNA.

On the other hand, domains II, III and especially V and VI have been strongly conserved. Large portions of these domains were found to be stripped from the 50S subunits upon treatment with T1 RNase (42). Furthermore, several experimental results suggest that these domains belong to functional areas of the 50S subunits. a) These domains contain single-stranded sequences which are very sensitive to T1 RNase within the 50S subunit and which are therefore good candidates for binding ribosomal ligands. (b) with the exception of a m⁶A residue located in domain IV, all the modified nucleotides which have been identified in 23S RNA are located in domain II (3), V (4) and VI (6). It is tempting to ascribe a functional role to these post-transcriptionally modified RNA regions as the nucleotide sequence around the modified nucleotides is highly conserved throughout evolution. c) Kethoxal has been used to identify accessible single-stranded guanine residues in 50S subunits. Some of these sites (positions 2307, 2308 and 2458 in domain VI) have been shown to be protected from kethoxal modification by 30S subunits and so have been postulated to occur at the subunit interface (43). d) Domain VI contains the binding site for protein L1 (3), which has been shown to be at the subunit interface (44) and to be indispensable for the fixation of the aatRNA (45). e) The region 2282-2390 of *E.coli* 23S RNA was found associated with 5S RNA and proteins L5, L18 and L25 (46) and it has been proposed that 5S RNA serves to bind the aatRNA (47). f) Two chloramphenicol resistant mutants in yeast mit 21S RNA have been analyzed and found to contain single base-substitution at sites corresponding to the positions 2447 and 2503 of 23S RNA (domain VI) (15,2). g) Greenwell et al. (48) have used a puromycin analogue for affinity labeling of ribosomes presumably in the A site. This analogue reacts with 23S RNA at a sequence reported to be either G-U-C^{*}-C-Gp or G-U-U-C-Gp. Although, such sequences are found at 4 positions in 23S RNA, only that at position 2553-2557 (domain VI) is single-stranded in our model. It is very accessible in the 50S subunit as it is cleaved by S1 nuclease. Therefore U2555 is probably the nucleotide cross-linked with the puromycin derivative. Points d, e, f and g strongly suggest that domain VI belongs to the A site.

Finally, since the genes for some large eukaryotic rRNAs contain introns it is worth noting that all intron sequences reported up to now occur in regions corresponding either to domain V or to domain VI of 23S RNA (Fig. 5). The rRNAs containing inserted sequences in such strategic domains probably lead to inactive ribosomes so that the process of excision and ligation of the intron sequences might regulate the level of protein synthesis.

ACKNOWLEDGMENTS

We are grateful to Dr. S. Vassilenko for his generous gift of *Naja oxiana* RNase. The excellent technical assistance of B. Muller and E. Schieffmayer is acknowledged. M. Schneider is thanked for her participation in the drawing. This work was supported by grants from the "Centre National de la Recherche Scientifique" and the "Délégation Générale à la Recherche Scientifique et Technique" and by grants from the Deutsche Forschungsgemeinschaft (SFB46) to H.K..

*To whom all correspondence should be sent.

REFERENCES

1. Branlant, C., Krol, A., Machatt, M.A. and Ebel, J.P. (1979) *FEBS Lett.* **107**, 177-181.
2. Brosius, J., Dull, T.J. and Noller, H.F. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 201-204.
3. Branlant, C., Krol, A., Machatt, A. and Ebel, J.P. (1981) *Nucl. Acids Res.* **9**, 293-307.
4. Machatt, A., Ebel, J.P. and Branlant, C. (1981) *Nucl. Acids Res.* **9**, 1533-1549.
5. Fox, G.E. and Woese, C.R. (1975) *Nature* **256**, 505-507.
6. Woese, C.R., Magrum, L.J., Gupta, R., Siegel, R.B., Stahl, D.A., Kop, J., Crawford, N., Brosius, J., Guttel, R., Hogan, J.J. and Noller, H.F. (1980) *Nucl. Acids Res.* **8**, 2275-2293.
7. Glotz, C. and Brimacombe, R. (1980) *Nucl. Acids Res.* **8**, 2377-2395.
8. Stiegler, P., Carbon, P., Zuker, M., Ebel, J.P. and Ehresman, C. (1981) *Nucl. Acids Res.* **9**, 2153-2171.
9. Edwards, K. and Kössel, H. (1981) *Nucl. Acids Res.* **9**, 2853-2869.
10. Van Etten, R., Walberg, M. and Clayton, D. (1980) *Cell* **22**, 157-170.
11. Eperon, I., Anderson, S. and Nierlich, D. (1980) *Nature* **286**, 460-466.
12. Wild, M. and Sommer, R. (1980) *Nature* **283**, 693-694.
13. Gourse, R. and Gerbi, S. (1980) *Nucl. Acids Res.* **8**, 3623-3637.
14. Rae, P., Kohorn, B. and Wade, R. (1980) *Nucl. Acids Res.* **8**, 3491-3504.
15. Dujon, B. (1980) *Cell* **20**, 185-197.
16. Allet, B. and Rochaix, J. (1979) *Cell* **18**, 55-60.
17. Nomiyama, H., Sakaki, Y. and Takagi, Y. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 1376-1380.
18. Hall, L. and Maden, E. (1980) *Nucl. Acids Res.* **8**, 5993-5996.
19. Bayev, A., Georgiev, O., Hadjiolov, A., Nikolaev, N., Zkryabin, K. and Zakharyev, V. (1981) *Nucl. Acids Res.* **9**, 789-795.
20. Branlant, C., Sriwidada, J., Krol, A. and Ebel, J.P. (1977) *Nucl. Acids Res.* **4**, 4323-4346.
21. Vassilenko, S. and Bakina, V. (1965) *Biokhimiya* **30**, 705-712.
22. Peattie, D. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 1760-1764.
23. Pipas, J. and Mc Mahon, J. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 2017-2022.
24. Vassilenko, S. and Rytte, V. (1975) *Biokhimiya* **40**, 578-582.
25. Branlant, C., Krol, A., Ebel, J.P., Gallinaro, H., Lazar, E. and Jacob, M. (1981) *Nucl. Acids Res.* **9**, 841-858.
26. Allet, B. and Spahr, P. (1971) *Eur. J. Biochem.* **19**, 250-256.
27. Morgan, E., Ikemura, T. and Nomura, M. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 2710-2714.
28. Fellner, P. and Sanger, F. (1968) *Nature* **219**, 236-238.
29. Sriwidada, J., Branlant, C. and Ebel, J.P. (1979) *Biochimie* **61**, 869-876.
30. Baer, R. and Dubin, D. (1981) *Nucl. Acids Res.* **9**, 323-337.
31. Branlant, C., Sriwidada, J., Krol, A. and Ebel, J.P. (1977) *Eur. J. Biochem.* **74**, 155-170.

32. Branlant, C., Krol, A., Sriwidada, J., Ebel, J.P., Sloof, P. and Garrett, R. (1975) *FEBS Lett.* 52, 195-201.
33. Sloof, P., Hunter, J., Garrett, R. and Branlant, C. (1978) *Nucl. Acids Res.* 5, 3503-3513.
34. Krol, A., Machatt, M.A., Branlant, C. and Ebel, J.P. (1978) *Nucl. Acids Res.* 5, 4933-4947.
35. Krol, A., Ebel, J.P. and Branlant, C., in preparation.
36. Edwards, K., Bredbrook, J., Dyer, T. and Kössel, H. (1981) *Biochemistry International* 2, 533-539.
37. Nazar, R. (1980) *FEBS Lett.* 119, 212-214.
38. Pace, N., Walker, T. and Schroeder, E. (1977) *Biochemistry* 16, 5321-5328.
39. Kelly, j. And Cox, R. (1981) *Nucl. Acids Res.* 9, 1111-1121.
40. Köchel, H., Lazarus, C., Basak, N. and Künzel, H. (1981) *Cell.* 23, 625-633.
41. Shatsky, I., Evstafieva, A., Bystrova, T., Bogdanov, A. and Vasiliev, V. (1980) *FEBS Lett.* 122, 251- 255.
42. Branlant, C., Krol, A., Sriwidada, J. and Ebel, J.P. (1977) *J. Mol. Biol.* 116, 443-467.
43. Herr, W. and Noller, H. (1979) *Cell* 18, 55-60.
44. Morrison, C., Garrett, R., Zeichhardt, H. and Stöffler, G. (1973) *Mol. Gen. Genet.* 127, 359-368.
45. Kazemie, M. (1976) *Eur. J. Biochem.* 67, 373-378.
46. Branlant, C., Krol, A., Sriwidada, J. and Brimacombe, R. (1976) *Eur. J. Biochem.* 70, 483-492.
47. Dohme, F. and Nierhaus, K. (1976) *Proc. Natl. Acad. Sci. USA* 73, 2221-2225.
48. Greenwell, P., Harris, R. and Symons, R. (1974) *Eur. J. Biochem.* 49, 539-554.