
Secondary structure model for 23S ribosomal RNA

Harry F.Noller, JoAnn Kop, Virginia Wheaton, Jürgen Brosius, Robin R.Gutell, Alexei M.Kopylov, Ferdinand Dohme and Winship Herr
Thimann Laboratories, University of California, Santa Cruz, CA 95064, and

David A.Stahl, Ramesh Gupta and Carl R.Woese
Department of Genetics and Development, University of Illinois, Urbana, IL 61801, USA

Received 27 July 1981

ABSTRACT

A secondary structure model for 23S ribosomal RNA has been constructed on the basis of comparative sequence data, including the complete sequences from *E. coli*, *Bacillus stearothermophilis*, human and mouse mitochondria and several partial sequences. The model has been tested extensively with single strand-specific chemical and enzymatic probes. Long range base-paired interactions organize the molecule into six major structural domains containing over 100 individual helices in all. Regions containing the sites of interaction with several ribosomal proteins and 5S RNA have been located. Segments of the 23S RNA structure corresponding to eucaryotic 5.8S and 2S RNA have been identified, and base paired interactions in the model suggest how they are attached to 28S RNA. Functionally important regions, including possible sites of contact with 30S ribosomal subunits, the peptidyl transferase center and locations of intervening sequences in various organisms are discussed. Models for molecular 'switching' of RNA molecules based on coaxial stacking of helices are presented, including a scheme for tRNA-23S RNA interaction.

INTRODUCTION

The last few years have seen a rapid development of the study of the large ribosomal RNAs (1, 2). In the main this reflects the introduction of rapid nucleic acid sequencing technology (3, 4) and the power of comparative sequence analysis in deducing secondary structure (5, 6). Although complete elucidation of the role of ribosomal RNA in ribosome function and assembly will doubtless require considerable three dimensional structural information, our present level of understanding of the 16S RNA structure has already provided significant insight into several aspects of ribosome biology (1). In this paper we present a model for the secondary structure of 23S ribosomal RNA. As in the case of 16S rRNA (6) evidence for the correctness of the model comes largely from comparative sequence analysis. The latter is based mainly on the nucleotide sequences of the *E. coli* (7) and *Bacillus stearothermophilus* (8) 23S rRNA genes. The two organisms represent the phylogenetic extremes of the eubacteria; their 23S rRNA sequences differ in 26% of their analogous positions. These data are supplemented by sequences of the corresponding

large subunit rRNAs from mammalian mitochondria (9, 10) and other partial sequences from the literature. Sites accessible to single strand-specific chemical and enzymatic probes are given as further evidence for our model. The 23S rRNA molecule is organized by long-range base paired interactions into six major structural domains, and exhibits many of the same kinds of helical structures seen in 16S rRNA.

METHODS

23S rRNA gene sequences

The E. coli 23S rRNA sequence was reported by Brosius et al. (7). A 23S rRNA gene from B. stearothermophilus strain 1054 was cloned in pBR 313 and pBR 322 and sequenced (8) by the method of Maxam and Gilbert (4). The two sequences were aligned for maximum homology. Additionally, the sequences of the human and mouse mitochondrial large subunit rRNAs were used (9, 10) as well as numerous partial sequences (11-19). After initial completion of our studies, the nucleotide sequence of the maize chloroplast 23S RNA gene became available (66). Use of this additional information resulted in changes in ten of the helices in our earlier model.

Secondary structure strategy

All nucleotide differences between the two aligned sequences were marked according to whether they were transitions or transversions. Using such notation on the aligned sequences, one can readily detect base paired regions common to the two RNAs that differ significantly in sequence. Thus the derivation of the secondary structure begins with those helices for which there is the strongest comparative evidence. In this way the number of potential helices in the molecule (which number is enormous) could be reliably reduced, greatly simplifying the task of deciding among the remaining helical possibilities, for which less or no comparative evidence existed. At the same time, fitting of the mammalian mitochondrial and other sequences to the developing structure brought additional data to bear on the remaining unstructured part of the sequence. [In cases where no base replacements are found between the two bacterial sequences, the more highly diverged sequences often show differences.] Computer-generated arrays (34) were used as a source of potential helices. As with the 16S rRNA (6) only helices containing four or more base pairs were compiled; this amounts to about 4×10^5 possibilities, of which only about 100 are the true helices.

Chemical and Enzymatic Probes

Bisulfite was used as a probe of single stranded cytosines in naked 23S

RNA and in 50S subunits (26). Kethoxal was used as a probe of single-stranded guanines in 50S subunits (24).

Naked 23S RNA was also probed with RNase T_1 under mild conditions (1:150 ratio of T_1 :RNA, 10 mM $MgCl_2$, 100 mM NH_4Cl , 10 mM Tris pH 7.2 for 30 min at 0°C. The reaction was stopped by addition of diethylpyrocarbonate (0.1%) and SDS (0.5%). Products were 3' labeled with ^{32}P Cp (21), fractionated by two dimensional gel electrophoresis (22) and sequenced by the chemical cleavage method of Peattie (23). In most cases 5' and 3' termini could be determined, giving the precise positions of RNase T_1 attack.

RESULTS AND DISCUSSION

Table I summarizes the comparative sequence evidence used in deriving the secondary structure model. In most cases, the E. coli and B. stearothermophilus sequences in the vicinity of each helix are shown. In many cases, the mammalian mitochondrial and maize chloroplast large subunit rRNA sequences (9, 10) are shown; and other partial sequence data (11-19, 64) are also included where applicable.

We consider the existence of a helix proven if there are two or more base pair replacements to support it. If only a single base pair replacement is known, we consider the helix likely, but not proven. [If a helix can be extended by bulging one or more residues, proof must apply to both sides of the bulge; however, if there are only 1-3 pairs on one side of a single residue bulge, a single base pair replacement is considered proof of that part of the composite structure.] In all other cases where helices are presented, the sequences are identical in all available organisms, and so the structure shown is merely consistent, not proven.

Agreement with the chemical and enzymatic probe experiments is generally good, but some apparent conflicts exist. Of the 30 guanine sites in 23S rRNA that are kethoxal-reactive in 50S subunits (and which can be located in the primary structure) (24), all but three are unpaired, in agreement with the model (Fig. 1). Two of the three conflicting sites occur near each other at positions 2093 and 2102, in a well-characterized helical stem implicated in the binding of protein L1 (see below). Thus there is the suggestion that this stem may be unpaired to some degree in the ribosome, although it appears to be highly resistant to RNase T_1 in the L1-23S RNA complex (13). Kethoxal reactivity of G_{2093} and G_{2102} is similar in 50S and 70S ribosomes (25).

Bisulfite-reactive sites in naked 23S RNA and 50S ribosomal subunits have also been studied (26). Of the 75 of these sites reliably placed, 64 occur in

Table I: Comparative sequence evidence for base pairing at specific sites in 23S rRNA^a

1.	5 GGUUAAGC.....GCUVAACC GGUUCNA.....UUUCNAC-	2900 Ec, Bs Zm chl	8.	140 GUGUGUU—CGACACAC CUGCCGUAAUGGGCCAG CGGCCCCGGGUUC—CUCGGGGCCACGC	Ec Bs Xl 5.8s
2.	20 GCGUACACGGU.....ACCGUGUACGU GGGCGACGGU.....ACCGUGUGCU	520 Ec Bs	9.	160 UCAUUAACUGNAUCCAUAGGUUAUGA UCCAUGCCUGAAUCCAUAAGGGCAUGA (5.8S)UGUCUGGG.....ACCUCAGUC(28S)	Ec Bs Xl 5.8S/28S
3.	40 UGGCAUGCAG.....CUGACUGACCG UGGCACUAGG.....CUAUGACCG	440 Ec Bs	10.	190 210 ACCGGGGNA.....ACCCGGGNA ACCGGGNA.....ACCCGGGNA ACUUGUGNA.....AGCCGAAAG ACCGCCGNA.....AAGCGGNA	Ec Bs Eg chl Xl 28S
4.	60 CGUSCUAAUCUGCG DGGGACACGCCGA CGCAGCCUAGTUGCG CGCAGC—AAACUGUG	Ec Bs Xl 5.8S Dm 5.8S	11.	235 260 GAUUCGCC.....CGGGGNC GAUUCGCG.....ACGGNAC GAUUUCUU.....CGGAAUC GAUUCGCGU.....AUGGAGC	Ec Bs Eg chl Zm chl
5.	80 CGUCGGU.....ACCGCG CUCGCGG.....DCCGGAG CUUWAG.....UCUWAG	110 Ec Bs Eg chl	12.	275 365 GAGC.....GCUUC AAAC.....GCUUC AAAC.....UUUU	Ec Bs Eg chl
6.	90 AAGGUGUAUGAACCGU GAGCUGUAAGCAAGCGU	Ec Bs	13.	290 350 CAGUGUGUGUAGUG.....CACAAAUGCCAUCCUG CAGUGAACGGGUAAGC.....GCGAAACUCCGUCUC	Ec Bs
7.	130 UGGGNAACCCA UGGGGNAACCCA UAGGUCAACCUU	Ec Bs Zm chl			

14. GCUCUGGAAAGCGCG EC 310
CCGGUCGNA⁵⁸⁵CGCCCG Bs
GUACAGC 1260 EC
ACCGCG Bs
AUUAAC Hs mt
15. GGUGACAGCCCC EC 600
AGGUGACAGCCCU Bs
UGGCUAUA⁶⁰⁰GCCA Zm ch1
UUUAUUCUGUAGC 660 EC
UGACGAUGCCGUCG Bs
GUACGUCGUCGUCG Bs
16. GAGUA 435 EC, Bs, Zm ch1
UACUC
17. GGGGGGA 395 EC
CGCGGGA Bs
GCAUGGGG Zm ch1
UCCCGUCU EC
UCCCGUCG Bs
UCCCGUGU Zm ch1
18. UGGGGGACCAUCCUCA EC 410
AGGAGGACCAUCCCCA Bs
AGCAAGGACCA⁶⁸⁰CUUGCA Zm ch1
19. CCCCGGCGAGGGG EC 490
CCCCGGAGGGGA Bs
ACCCCA⁶⁸⁰GUGGU Zm ch1
20. UACAA⁵³⁰G 570 EC, Bs
UUUGUA
21. GUGG 560 EC
GUCA Bs
UUGC
22. CACGCUAGCGGUG EC 550
CGGUCUAGCGUG Bs
GGGMAAGUGAUCUCU Zm ch1
23. GUACAGC 585 EC
ACCGCG Bs
AUUAAC Hs mt
24. UUUAUUCUGUAGC 600 EC
UGACGAUGCCGUCG Bs
GUACGUCGUCGUCG Bs
25. GGUUAACCG 620 EC
GGUUAAGCC Bs
GGUUAAGG Zm ch1
AGGGAAACCGAGUCUU EC 630
CAGCGAAAGCGAGUCUG Bs
26. ACCGGGU 800 EC
ACCAAGGA Bs
CCUGGGU Zm ch1
ACCAGAC Hs mt
ACCAAC Mm mt
27. ACCGGGU 800 EC
ACCAAGGA Bs
CCUGGGU Zm ch1
ACCAGAC Hs mt
ACCAAC Mm mt
28. UCUAGCCAU⁶⁹⁰GGG 770 EC
UCU⁶⁹⁰CCCAUGC Bs
UCU⁶⁹⁰CCCAUGAC Zm ch1
GCU⁶⁹⁰ACCUAAGAA Hs mt
GCU⁶⁹⁰ACCUA⁶⁹⁰AAA Mm mt
- 28a. UAGCC 700 EC
UACCC Bs
UAUCC Zm ch1

29.	705 <u>GUUU</u> <u>GACC</u> GGGU..... <u>GCCC</u> GGAU..... <u>GUCC</u> GCUA..... <u>GAGC</u> AUUU..... <u>GAUU</u>	730 Ec Bs Zm chl Hs mt Mm mt	870 UGUUUGG..... <u>CCGAUCCA</u> UGAUUGG..... <u>CCCAGUCA</u> UGUUUGG..... <u>UCGAGGCA</u>	905 Ec Bs Zm chl
30.	710 GGUUGG..... <u>CUAAAU</u> GGCCGG..... <u>CCGGCU</u> GCUUGG..... <u>CUAAGG</u>	720 Ec Bs Zm chl	885 GGGGUC..... <u>GACUUACC</u> GGGGCC..... <u>GGUUUACC</u>	895 Ec Bs
31.	740 CCGACUAAU..... <u>AUUAGCGG</u> CCCACGCAC..... <u>GUCGGGG</u>	760 Ec Bs	950 <u>CACGGCGGGUUAACGUCGUGUG</u> <u>GACUCGAGUGAUAAGAUCCGUGGUC</u> <u>GACGAUGGGGAUAAGCUUCAUCGUC</u>	Ec Bs Zm chl
32.	815 CUCCCC..... <u>GGGGAG</u> GUCCAA..... <u>UUGGAC</u> ACCCAA..... <u>UUGGGU</u>	1190 Ec, Bs Hs mt Mm mt	995 <u>CCGCCAGCU</u> <u>AGCUCGG</u> <u>UCGCCAGCU</u> <u>AGCUCGG</u> <u>UCACCAGCU</u> <u>AGCUCUGG</u>	1160 Ec, Bs Bs Zm chl
33.	830 AAAGCUUUUAGUAGGCC AAAUAGCUUUAGGGCUAGCC AAAUAGUUUUAAGUUCAAUU	Ec Bs Mm mt	1005 AGGUC..... <u>CACCG</u> AGGC..... <u>CGCCG</u>	1150 Ec Bs
34.	845 CUCGUGA..... <u>UCACGGG</u> CUCGGG..... <u>CCCGGG</u> GUUGACU..... <u>AGUCGGC</u>	940 Ec Bs Zm chl	1015 CAUGGU..... <u>AACCAUG</u> GCACGU..... <u>AAGGUC</u>	1145 Ec Bs
35.	850 CAUCUCCGG..... <u>CCGGAGAUG</u> GAGUUGUG..... <u>CCAACGACUU</u>	930 Ec Bs	1040 CGAUGUGGAAGGC..... <u>GUCGGCCUGCG</u> GGAUGUGGAGUUG..... <u>GUGACUCUGCG</u> GGAGTUGGGGGUG..... <u>GCGCCCUUGCG</u>	1120 Ec Bs Zm chl
40.			1055 GCCAG..... <u>CUGGU</u> ACCAG..... <u>CUGGU</u>	1105 Ec Bs

- 1060
 44. GAUGTUGGCUUAGAACGCGCCAUCAUU Ec
GAUGTUGGCUUAGAACGCGCCACCAUU B8
GAGGUUUGCCUAGAACGCGCCACCCUU Zm chl
- 1170
 45. CGCGCAGCGCCGUUAUGCCUGUUGGGUA Ec
GCGG--GAUGACCGUU--GGUCAUC--GGUA B8
GUGG--GAUUCMAAU--GCAUC--GGUA Zm chl
- 1210
 46. UUCUUAAGCCUCGCGAAGGUGUGUGGCAUGCUGGAGGUAUCAGAA Ec
UUCUUA--GGCGGUAGAACGCCAGCCGGAAAGACUGGUGGAGCGCUUAGAA B8
- 1240
 47. AGCGGGUUGAAGCCCGCU Ec
CAGAGGUGAUAUCCUCUG B8
- 1290
 48. 1300 1645
CGCC.....GGCG Ec, B8
GGGG.....CCCC Zm chl
UUCU.....AGGA H8 mt, Mm mt
- 1305 1630
 49. GACCA.....UGGU Ec
GCCUA.....UAGC B8
ACCCA.....UGGU Zm chl
- 1315 1335
 50. UCCUGUCCAAG.....CGGGCAGGG Ec
UCCUGAAGAG.....CCGUCAGGG B8
UCCUCCGCAAG.....CCACGGGG Zm chl
- 1380
 51. AGUCG AUGGAA Ec
AGGUAUGGACAA B8
- 1390 1400
 52. ACAGG.....CCUGU Ec, B8, Zm chl
- 1410 1590
 53. UUGGUG--UUAUC.....GGUACAUCAA Ec
CACCUCUUCUCC.....GGGAGAGGUG B8
- 1425 1575
 54. AGGGG.....CCUCU Ec
UGGGG.....CCUCU B8
- 1435 1560
 55. GGAGNAGGC.....GCCUCUCUC Ec
GCAGGAGGA.....UUCACACUGC B8
- 1445 1550
 56. UGUUG.....CAACA Ec
GGCG.....UCCCC B8
GGTUA.....UAACC Zm chl
- 1450 1465
 57. CCGGG.....CCCGG Ec
CAGGC.....GCCUG B8
- 1470 1520
 58. UUAAGCGUG.....GGCGUAUGA Ec
CCAAAGCCGC.....GCGGUAUGG B8
UAGAGCAC.....GGUCCGAGUA Zm chl
- 1500
 59. GGCUGUUUCCAGGCMAAUCGGMAAU--CAAGGCU Ec
GGCU--GAUCCGUAGGCMAAUCGGGUAUUGAAGCC B8

60. 1540 GAGGCACUACGGGUCUGAA Ec
 GACGGAGUCUCGUCCGGAA Bs
 CAGGCCUACGGGC-GAA Zm chl
61. 1640 CAGGUAGAGAAUACCAAG Ec
 CGNGGAGAAUCCUAG Bs
 UAGGUAGAGAAUACCUAG Zm chl
62. 1650 2010 UGNAGAA.....ACUCAGUG Ec
 CCGGAGGA.....ACCCGGUG Bs
63. 1660 2000 CUCGGUG.....CACCCGAG Ec
 CUCUCGU.....ACGAGAG Bs
64. 1685 1705 GUGCC.....GGCAC Ec
 ACCCC.....GGGGU Bs
65. 1710 GCTGUAUUGAGGUGAGGUCCUCGCGGAGGACUGAAUACAGU Ec
 GCUGUUGGGUGAAGACCCGAGCAGC Bs
66. 1770 1985 CUGGCCGA.....UGNUGCCAG Ec
 GCCCAAGCG.....CGACUUGGC Bs
67. 1780 UGUUUYAANAACA Ec, Bs, Mm mt, Hs mt
68. 1795 1820 GCACUG.....UACGGUGU Ec
 GGUCUCUG.....UAGGGCCU Bs
69. 1910 ACACGAAUGG Ec
 AGUCGUAAGCG Bs
70. 1835 1975 ACGCC.....GGCGU Ec
 ACACC.....GGUGU Bs
71. 1845 1895 UGCCGG.....CCGGUA Ec
 UGCUGG.....CCAGUA Bs
72. 1865 1880 GGGGU.....AGCUCU Ec
 GCGCU.....AGGUGC Bs
 GUUGU.....GCCGCG Zm chl
73. 1910[†] GCCGYAACUA[†]YACGGUC Ec, Bs
 CGCGGU-AUCCUGACCGUG Mm mt
 GGGGUAACUAUGACUCUC Xl 288
74. 1940 UAGGUGCGAAATUCCUUG Ec, Bs
75. 1950 GUCGGUAAAGUCCGAC Ec, Bs, Zm chl
 GUUCUUAU[†]UAGGAC Mm mt
 CGUCACUAAUUA[†]GUGCG Xl 288
76. 2040 ACUCGUGUAGAAUGCAGUGA Ec
 ACUACUGUAGAAUGCAGGUA Bs
 CCUUCAGGAAAGAGGCTUGAAU Mm mt
 AGCCAGGNAACGGGCTUUGGG Xl 288

- 2050
 77. CCCGCGCAAGCGG.....CCUAUCGCCGUGGG EC
CCCGCAGAGACGG.....CCUAUCCGUGCGGGG B8
CCUCACCCUGACAG.....CAUAUCCGUGUGGG Zm chl
AUAUAUAUAAGCGA.....UCUAUCUAUUUAACGA Mm mt
- 2445
 78. CCCGUGA.....UCCGGGG Ec, B8
CCCUAUGG.....CCUAGGG Mm mt
CCUUGUG.....CACAGGG Xl 288
- 2240 2435
 79. ACUUUAUAU.....GUAGUU.....AAAGGU EC
AGUUUAACUG.....GCAGUU.....AAAGCU B8
AGUUAAUUU.....AAUUU.....CAAGUU Mm mt
AGUUAGACUU.....GGAGUU.....AAAGUU Xl 288
- 2090 2230
 80. GCUUGACACU.....AGUGUCUGU EC
GCCUGUAUG.....AGUGUCAGC B8
AACUUAUCUA.....UAGUUAUAGUU Mm mt
- 2100 2190
 81. CUGAACAUUGAGCCUUGA.....UUAAUGUUUGAUUCUA EC
UGAAUUUUUGUGUCGU.....GGCGAUUUUGAAUUCUA B8
GUGNAGACAUAGAGG.....ACUCUUAUCGUUUUUCAC Xl 288
GCAUAGUAGCCUUAAGG.....CCUUUUGCCGUUAUUUGC Dd
- 2125 2175
 82. GGUGG.....CCACC Ec, B8, Dd
AGUGG.....CCACU Xl 288
- 2130 2160
 83. GCCUUU.....GGAGCC EC
GCCUGG.....GGAGCC B8
- 2150
 84. GUUGAGCGCCAGUCUAGU EC
GCCGAGCGCCAGCUUCGU B8
GGCCCCUUCGGGGGGCC Zm chl
- 2210
 85. GACCCGUAUCCGGGUU EC
CACCCCUAUGCAGGGUG B8
- 2250
 86. GACUGGGGGGUC Ec, B8
CUAUGGGGCGUAG Zm chl
GSUUGGGCGACC Hs mt
- 2260
 87. CUCCUCUAAAGAGUACGGAGGAG EC
CGCCUCCCAAAAAGUAAACGGAGCGG B8
CUCGGAGCAGAACCCACCTUCGGAG Hs mt
- 2290 2340
 88. AGGUUGGU.....AGCCAGCUU EC
AGGUUCCU.....AGGGAGCUU B8
- 2305
 89. UCCUGGCGGACAUACAGA EC
GAAUGGUUGAAUAUCAUUC B8
- 2330
 90. AGUCCAAUGSCAUAA EC
AGUCFAAAGGCACAA B8
AAUAUAUUGUAUAA Sc mt
- 2350
 91. CUGCGAGCGUGAGCGGCGGAGCAG EC
CUGCGAGCGGACAGGUCGAGCAG B8
UDGUAAACAAACGAGUGUAAACA Sc mt

- 2380
 92. GUGGAAAGCAG EC
GGACGAAAGUCG B8
GUACGUAAGUAV SC mt
- 2400
 93. CGGUGUUCUGAUAUGAAGGCCCAUCG EC
CGGUGUUCGCAUGAAGGCCCAUCG B8
AAUAACAUGAUGUAAGGUUAUUG SC mt
- 2460 2490
 94. GCTUGAUAACGCC.....GGCGUGUUGCC EC
GCUGAUCUCCCC.....GGGAGGDUUGCC B8
CGCAUCCUAUU.....AUAAGGDUUACG B8 mt
GGUAUAUAACG.....AGUUAUGDUUGCC SC mt
- 2510 2580
 95. UCGGCUC.....GAGCUGG EC, B8, Zm ch1
UUGAUC.....GAUCUGA H8 mt, Mm mt
UCGACTC.....GAGTUGG SC mt
- 2515 2570
 96. AUCAC.....GUGGU EC
AUCGC.....GCGGU B8
AGGAC.....GUCUU Mm mt
AUCAU.....AUGUU SC mt
- 2530
 97. AUCCUGGGCUGAUAUGUCCCAAGGU EC
AUCCUGGGCUGUAUCGUCUCCCAAGGU B8
AUCCCAUGGUGUAAGCUUAUUAUGGU Mm mt
UCCUUGGUGUAAAGCUAAGAAAGGU SC mt
- 2550
 98. AUGCGUUCGCCAU EC
UGGCUGUUCGCCA B8, Zm ch1
UCGUUUGUUAACGA Mm mt
UUGACUGUUCGUCAA SC mt
- 2600
 99. GNACGUCGAGACAGUUC EC, B8
GACCGAGYAAUCCAGGUC H8 mt, Mm mt
AUACGAUGUAGAUACAGUAV SC mt
AAACGUCGUGAGACAGUUV Cr ch1
- 2635 2785
 100. GAGAAACU.....AGUUCUCC EC
GGAAAUUU.....AGAUUUCC B8
- 2645 2770
 101. GGGG.....CCCC EC
GAGG.....CCUC B8
- 2650 2670
 102. CUGCUCC.....GGAGUGG EC
CUGUCCU.....GGGGUGG B8
CCUUUCU.....UGAAAGG Cr ch1
- 2680 2730
 103. CGCAU.....AUGCG EC
CGCAC.....GUGCG B8
- 2700
 104. CGGGUUCUACUCCAAUGGCAC-UGCCCG EC
CCAGUUCUCCGUAAGGGCAC-CGCUUG B8
CCAGUUDCGUGCCUACCGUAAACGCUUG Zm ch1
- 2745 2760
 105. GUGCU.....AGCAC EC
GCGCU.....AGCGU B8
ACUUC.....GUAGU Zm ch1
- 2800
 106. UGACCCUUUAGGGUCC EC
CACCGUCAGCGGGU B8
UGCAAUGUCAUUUUGCG SC 288
CACGUUCGUGGGAAAC X1 288

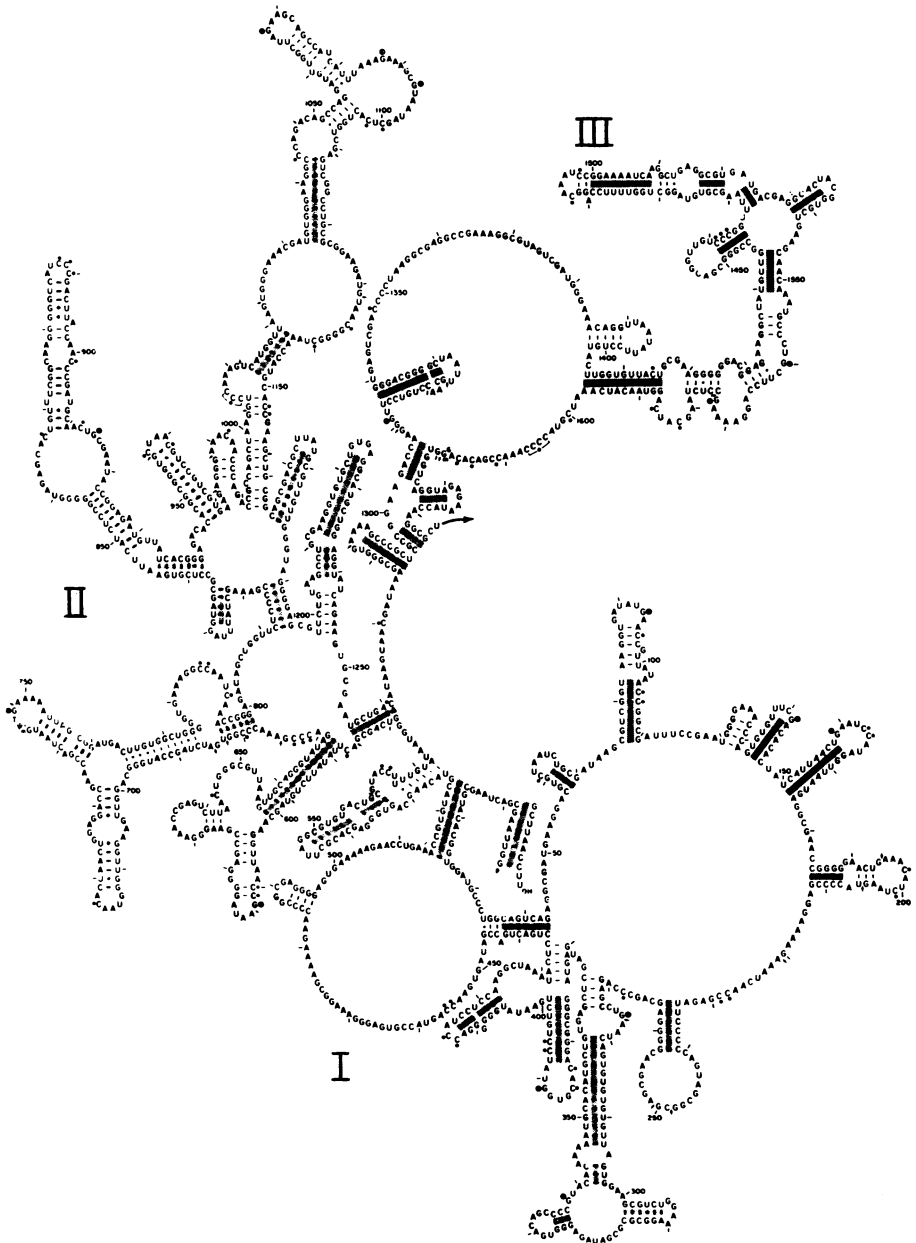
	2820	
107.	<u>AACGUUGAAGACGACGACGUU</u>	Ec
	<u>UCCUCGGAAGAUGACGAGGUC</u>	Bs
	<u>UAGCGCGAGACGAGCCGUU</u>	Zm chl
	2840	2880
108.	<u>UAGGCCGG.....CCGGUACUA</u>	Ec
	<u>UAGGUCCG.....CGGAUACUA</u>	Bs
	<u>UAGGUGUC.....GGCAUCCUA</u>	Nt chl 4.5s
	2860	
109.	<u>AGCGCAGCGAUGC GUU</u>	Ec
	<u>AGCGUGGCGACACGUG</u>	Bs
	<u>AGUGCAGUGAUGUAUG</u>	Nt chl 4.5s

^a Base paired sequences are underlined. Sequences are designated Ec, *Escherichia coli* (7); Bs, *Bacillus stearothermophilus* (8); Eg chl, *Euglena gracilis* chloroplast (11); Xl, *Xenopus laevis* (12, 15); Nc, *Neurospora crassa* (64); Dd, *Dictyostelium discoides* (13); Hs mt, human mitochondria (9); Mm mt, mouse mitochondria (10); Sc mt, yeast mitochondria (16); Cl chl, *Chlamydomonas reinhardtii* chloroplast (18); Nt chl, tobacco chloroplast (19); Zm chl, *Zea mays* chloroplast (66). Due to space limitations, only representative comparative examples are shown for each helix.

single stranded regions, in agreement with the model. Of the ones that occur in double stranded regions (and so are not compatible with the model) all but the ones at positions 96, 1152 and 2683 are located in the terminal base pair of a proposed helix.

Free *E. coli* 23S RNA was treated under mild conditions with RNase T₁ and the resulting fragments end-labelled and sequenced by the chemical method of Peattie (23). In this case, the fragments can be placed unambiguously in the sequence because of their length, and their 5' and 3' termini can usually be identified precisely. Here, 51 out of the 59 identifiable T₁ cleavage sites are in agreement with the model. In most of the conflicting cases, the site of cleavage is at a G·U pair at the end of a helix (e.g. position 2083) or in thermodynamically unstable structures that could be stabilized by proteins in the intact ribosome (e.g., positions 205, 597, and 1002).

Thus, the probing experiments are in generally good agreement with our proposed model. Some disagreements remain however, for which we do not have clear explanations at present. In a few specific cases (e.g., the helices at positions 1002-1004/1151-1153, 2676-2680/2727-2731) the conflicting evidence suggests that we accept the proposed structure with caution; in other cases (e.g., 678-683/794-799, 1478-1492/1498-1513, 2323-2326/2331-2334), the comparative evidence is overwhelmingly in favor of the proposed structures.



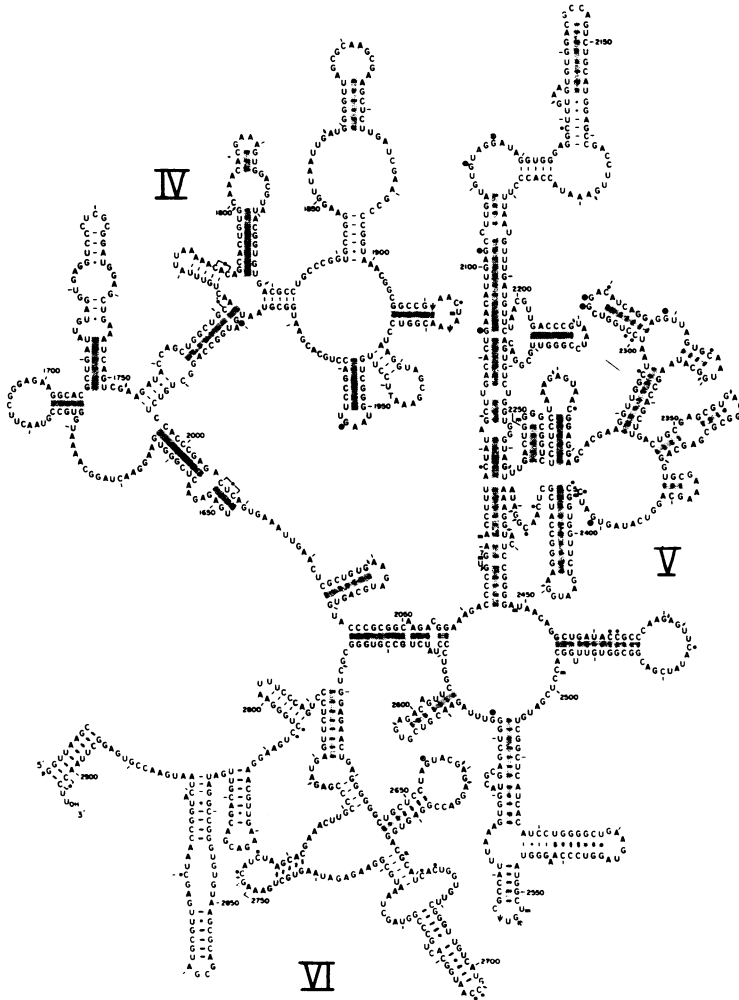


Figure 1. Secondary structure model for *E. coli* 23S ribosomal RNA. Primary structure is from Brosius *et al.* (7). The molecule is arbitrarily displayed in two halves (positions 1-1647 and 1648-2904); the pairing of the 5' and 3' ends is shown in both drawings. I-VI refer to the six major structural domains. Guanines reactive with kethoxal in 50S subunits (24 and W. H. and H.F.N., unpublished) are indicated by a circled K. Cytosines reactive with bisulfite in naked 23S RNA (26) are shown by a filled circle; unreactive cytosines are shown by open circles. Positions of cleavage of 23S RNA by RNase T₁ under mild conditions are indicated by arrows. Helices that we consider proven by comparative sequence criteria (see text) are stippled.

Here, we must conclude either that the true structure became disrupted under the experimental conditions (e.g., in the case of naked 23S RNA) or that multiple conformations are possible at these sites.

General architecture of the 23S ribosomal RNA

Our secondary structure model is displayed in Fig. 1. About 52% of the residues exist in paired structures. The 5' and 3' terminal sequences are base paired, giving the whole the form of a closed loop. As in the 16S RNA (1, 6), long range base paired interactions partition the chain into readily identifiable large domains. There are six of these in the 23S RNA, defined by their respective long range interactions: domain I (16-25/515-524), domain II (579-585/1255-1261), domain III (1295-1298/1642-1645), domain IV (1656-1664/1997-2004), domain V (2043-2054/2615-2625) and domain VI (2630-2637/2781-2788). These domains project from the central loop created by pairing of the 5' and 3' ends of the molecule.

The 23S RNA chain is readily cleaved into a 13S and an 18S fragment (20), and the 18S is further cleavable into an 8S and a 12S fragment (63). Comparison of oligonucleotide compositions of the various fragments with the complete 23S RNA sequence suggests that the 13S fragment corresponds roughly to domains I and II (although the 13S/18S cleavage site appears to occur within domain II), the 8S fragment to domains III and IV, and the 12S fragment to domains V and VI. No "knots" (27) appear in our model; however, it would be premature to rule out their existence in the structure. Nor is any convincing evidence available to suggest how the structure might be organized into the compact, roughly spherical shape demanded by the 50S ribosomal subunit.

Electron microscopic analysis of unfolded 50S subunits or partially denatured 23S RNA has shown characteristic structural features that can be compared with our secondary structure scheme. 50S subunits unfolded by removal of magnesium ions appear as asterisk-like shapes, with five prominent arms radiating from a common center (28). These could well correspond to the five largest domains (I-V; Fig. 1), which radiate from the central loop. Large loop structures have been seen in electron micrographs of partially denatured 23S RNA (29), presumably corresponding to strong long-range base paired interactions between approximately positions 1/500, 1700/2100 and 2200/2600. These agree well with the strong long range interactions enclosing domains I (20/520), IV (1650/2000) and V (2050/2620). Hairpin structures visualized at approximately positions 950, 1550 and 2800 probably correspond to the extended compound helical stems centered at positions 890, 1500 and 2700.

As observed previously in 16S RNA, there are no long, uninterrupted regu-

lar helices. Instead, the 23S rRNA architecture appears to comprise an arrangement of small helices, many of which contain characteristic irregularities. As in the 16S RNA case, the helices in 23S RNA appear to be of several types. Some are regular (involve only Watson-Crick base pairs). Others are notable for a high content of G-U pairs. Still others are even more irregular, containing non-Watson-Crick "pairings" (especially A-G "pairs") and bulged residues. In 16S RNA there appeared a marked tendency for highly conserved residues to be concentrated in the non-helical regions. This tendency is not so pronounced in the present case. What is noticeable in the present instance, however, is that some of the helices are particularly variable in sequence phylogenetically (again a property of some helices in 16S rRNA).

Sites of interaction with ribosomal proteins and 5S RNA

Considerable data have been reported on the ribosomal protein binding sites in 23S RNA, although generally not as extensive as for 16S RNA (reviewed in ref. 37). These data come either from nuclease protection studies, in which a specific RNA fragment is shielded by a bound protein or group of proteins, or from covalent crosslinking. The regions bound by proteins L1 (35), L20 (36), L23 (38), and L24 (31, 38) can be located within their protected fragments, and the position of L4 can be inferred from ultraviolet irradiation-induced crosslinking of this protein to position 615 of the 23S RNA (39). Approximate locations of these proteins in the 23S RNA secondary structures are shown in Figure 2.

An interesting and characteristic feature noted previously in 16S rRNA (1) and 5S rRNA (30) is a class of helix containing a single bulged nucleotide, usually an A. The occurrence of this type of structure in regions of rRNA known to contain recognition sites for ribosomal proteins has been noted and it has been suggested that the bulged base could form part of the recognition signal for certain proteins (1, 30). There are fifteen singly bulged nucleotides in our 23S RNA model; ten of these are adenylate residues. We suggest that the bulged adenylate at position 443 is involved in the recognition of protein L24 (31), and that other helices of this type may also be involved in protein binding.

The type of helix containing multiple G-U pairs may constitute another class of protein recognition sites. Examples of proteins whose binding sites contain this type of helix include S8 in 16S RNA (32) and L25 in 5S RNA (24). In 23S RNA a number of these G-U helices occur, including 588-601/656-669 [implicated in binding protein L4 (34)] and the pair 2093-2103/2186-2196 and 2127-2143/2148-2161 [both implicated in binding protein L1 (13, 35)].

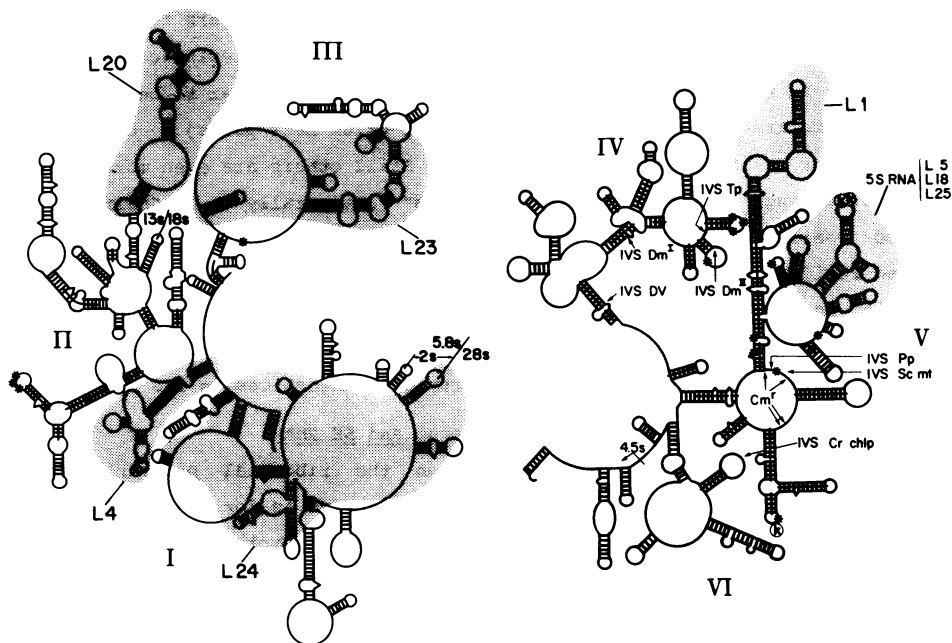


Figure 2. Schematic diagram of the 23S RNA secondary structure model. Regions containing binding sites for certain ribosomal proteins and 5S RNA are shaded. Helices that have clearly recognizable counterparts in mammalian mitochondrial large subunit RNAs are indicated by a dark bar. Post-transcriptionally modified nucleotides are shown by asterisks. I-VI are the six major structural domains. IVS, intervening sequence; Dm, *D. melanogaster*; Dv, *D. virilis*; Pp, *Physarum polycephalum*; Sc, *S. cerevisiae*; Cr, *Chlamydomonas reinhardtii*; Tp, *Tetrahymena pigmentosa*; mt, mitochondria; chlp, chloroplast; Cm^k indicates sites of nucleotide substitutions in chloramphenicol-resistant mitochondrial ribosomes (see text). Kethoxal-reactive sites protected in 70S ribosomes (25) are shown by a circled k. Positions corresponding to the points of demarcation of 5.8S, 2S and 4.5S RNAs are indicated.

RNA binding site fragments for proteins L20 and L23 are not well documented, but they can be placed approximately within the structure from the available data in the literature. Protein L20 is located in domain II and probably binds to the 23S RNA somewhere between positions 1000 and 1150 (36). Protein L23 binds in domain III and can be placed approximately between positions 1320 and 1600 (38).

The mechanism of association of 5S RNA with the 50S ribosomal subunit is not yet understood. Although it binds proteins L5, L18, L25 and L31' (33, 65), and requires L5 and L18 to associate with 23S RNA, it is not known whether or not it interacts directly with 23S RNA. Comparative sequence evi-

dence from B. stearothermophilus shows that the remarkable complementarity of base pairing observed between E. coli 5S and 23S RNAs involving position 143-154 of 23S RNA (40) is not maintained in general. Thus, it is unlikely that base pairing between 5S and 23S RNAs, although seemingly possible, actually occurs at this site. Protein-mediated binding between the two RNAs is, of course, a possibility.

The site of interaction of the complex between 5S RNA, L5, L18 and L25 with 23S RNA has been studied by the nuclease protection approach in two laboratories (41, 42). Both groups isolated a protected fragment of 23S RNA spanning approximately positions 2250-2350. Significantly, this region of 23S RNA is deleted in the mammalian mitochondrial large subunit RNA, the ribosomes of which lack 5S RNA.

5.8S, 2S and 4.5S rRNAs

In eucaryotic cytoplasmic 60S ribosomal subunits, a 5.8S RNA is found, in addition to a 26S (or 28S) and a 5S RNA. Although previously thought to be a 5S RNA analogue (43) more recent evidence suggests that 5.8S RNA is in fact analogous to the 5' terminus of bacterial 23S RNA (44), whereas the eucaryotic 5S RNA is at least structurally a close analogue of procaryotic 5S RNA (33, 45). Alignment of 5.8S RNAs with 23S RNAs so as to maximize their sequence homology shows that both can form similar base paired structures within their common regions (see Table I). Attachment of 5.8S RNA to 28S RNA, known to involve the 3' terminus of 5.8S RNA (47), is accounted for by its pairing with the 5' terminus of 28S RNA, which is a close analogue of the helical stem 150-158/168-176 in the E. coli structure (Table I; Figure 1).

Similarly, 2S RNA, found in some insect 80S ribosomes (46, 48), corresponds approximately to positions 140 to 165 of 23S RNA, and is probably bound to 5.8S RNA by base pairing analogous to the 132-137/142-147 pairing of 23S RNA. The 4.5S RNA found in certain higher plant chloroplasts has previously been postulated to be analogous to the 3' terminus of 23S RNA (49). This is supported by the finding that these two molecules can also form closely analogous secondary structures involving their homologous regions (R. Hallick, personal communication; Table I). By analogy with 5.8S RNA, it is possible that binding of 4.5S RNA to 23S RNA occurs via base pairing of residues homologous to the helix 2791-2796/2800-2805 of E. coli.

Intervening sequences in large subunit ribosomal RNAs

In several, but certainly not all cases, intervening sequences have been discovered in the large subunit RNAs of eucaryotic cytoplasmic, mitochondrial and chloroplast ribosomes (16-18, 50-53). Positions of six intervening

sequences are shown in Figure 2 in the analogous positions in the *E. coli* 23S RNA structure. Inspection of the primary and secondary structure around these sites shows that (1) there is no obvious common primary sequence among them, (2) all of the sites occur in single stranded positions, although nearly all are closely adjacent to helical elements, (3) all are located in highly conserved structures in the 3' half of the molecule, which include the peptidyl transferase region (see below).

The role of intervening sequences in ribosomal RNA genes is not yet clear; with the possible exception of the *Drosophila virilis* site (51), they do not appear to demarcate structural domains, as has been suggested in the case of protein genes (54). A more striking correlate is their appearance in what would seem to be functionally indispensable regions of the structure. It is likely that a ribosomal RNA containing an intervening sequence at any of the observed positions would be non-functional. Perhaps, then, they are involved in the regulation of the size of the active ribosome population in certain cells. Alternatively, they might constitute a proof-reading mechanism for ribosome maturation: only those ribosomes which have assembled properly would be recognized and processed. In any event, the fact that they have all been localized thus far in the 3' half of the large subunit RNA is probably significant.

Functional Sites

There exists a growing body of evidence implicating the 23S RNA in ribosomal function. Affinity labelling experiments, in which chemically reactive groups attached to tRNA or antibiotics are allowed to react with ribosomes, have shown that some of these functionally important ligands bind in the immediate vicinity of the 23S RNA (55). In two cases, attempts have been made to localize the precise sites of reaction with 23S RNA. In the case of iodoacetyl-Phe-tRNA (56), an RNA sequence was reported which, however, is not found in the complete 23S RNA sequence. Another affinity reagent, 5'-O-(N-bromoacetyl p-aminophenylphosphoryl)-3'-N-L-phenylalanyl puromycin aminonucleoside, an analogue of the antibiotic puromycin, was localized to the sequence G^{*}CCG or GU^{*}CG (57). Further studies are required to establish the precise sites of attack by these reagents.

Chloramphenicol, a specific antibiotic inhibitor of peptidyl transferase, is thought to interfere with proper binding of the aminoacyl tRNA at its acceptor end (for a review, see ref. 58). Certain chloramphenicol-resistant mutant mitochondrial ribosomes have been shown to arise by mutation at sites in the large subunit ribosomal RNA (16, 59). The region of the RNA containing

the mutations is highly conserved phylogenetically, indicative of its importance, allowing the sites of mutation to be located unambiguously in the *E. coli* 23S RNA sequence, at positions 2447, 2451, 2503 and 2504. In the secondary structure model (Figs. 1, 2) these four sites are found clustered around a central loop in domain V. Thus, this highly conserved region of the 23S RNA structure is likely to constitute some part of the peptidyl transferase center of the ribosome.

Kethoxal has been used as a functional probe of the 23S RNA by virtue of its reduced reactivity toward specific sites in the presence of functional ligands. Binding of 30S ribosomal subunits decreases the kethoxal reactivity of G₂₃₀₇, G₂₃₀₈, and G₂₅₅₃ (25), located in domain V. The simplest interpretation, that these sites make contact with the 30S subunit across the subunit interface, further implicates domain V in ribosome function.

Comparison of the 23S RNA with its mammalian mitochondrial ribosome analogue (9, 10) is of interest, since the latter molecule is only about half the size of 23S RNA. Presumably, sequences and structures which are indispensable for ribosome function are retained in the mitochondrial RNA, while less critical portions of the molecule have been lost during evolution. Homology between the bacterial and mitochondrial sequences is readily detectable, and with careful alignment of the respective sequences, conservation of certain secondary structure elements is also evident (cf. Table 1). Those helices that have clearly recognizable counterparts in the mammalian mitochondrial large subunit RNAs are indicated by a dark stripe in Figure 2. Interestingly, nearly all of the conserved helices are found in domains II, IV and V. Coincidentally, nearly all of the post-transcriptionally modified nucleotides also occur in domains II, IV and V. Thus domains II and IV, in addition to the well-documented case of domain V, are also likely to contain functionally indispensable regions.

Higher order structure and mechanism in 23S ribosomal RNA

The ultimate picture of ribosomal RNA will show the various helical elements in specific three dimensional arrays, interacting with other components, and in all probability undergoing movement during the process of translation. Although the secondary structure and its present implications are a significant beginning, we are still a long way from such a conception of the ribosome. The secondary structures and elements defined by them suggest certain possibilities for higher order structure and mechanism (1, 2, 5). Of particular interest are coaxial helices, for such structures at very least constrain the overall shape of the RNA significantly. The precedent for considering

coaxial helices is, of course, the tRNA molecule, which contains two such juxtapositions—the coaxial T ψ C and CCA arms, and the nearly coaxial anticodon and D arms. Any two adjacent helices (i.e., having no unpaired residues between them in one of the strands) should be considered as potentially coaxial. There are about 13 such possible coaxial situations (not so drawn) in Fig. 1.

One type of comparative evidence strongly suggests coaxiality. Suppose two adjacent helices in an RNA of organism A are of lengths n and m pairs. If the analogous helices in organism B are of lengths $n+a$ and $m-a$ (i.e., the sum of pairs in the two, $n+m$, is constant) then the two helices may be coaxial. An example of such a case for the 23S rRNA is helix 79 in Table 1.

Going further, the forming and unforming of such coaxial structures, and the alternate forming of mutually exclusive coaxial structures, should be considered as a possible basis for mechanism in the ribosome. It is conceivable that such a mechanism operates in the tRNA molecule during translation and in the 5S RNA as well.

As an example of how coaxial stacking might be employed in translation, we consider the following. At the 3' terminus of all tRNAs is found the invariant sequence CCA, likely to play a role in the precise positioning of aminoacyl and peptidyl moieties in the peptidyl transferase site of the 50S ribosomal subunit. The two cytidine residues have been shown to be crucial for productive binding of oligonucleotide analogues of peptidyl tRNA to the peptidyl transferase site (60) (although the 3'-terminal adenosine of f-Met-tRNA can be replaced by inosine or guanine and still carry out peptidyl transfer [61]). Peattie and Herr (62) have demonstrated that the two 3'-terminal C residues in ribosome-bound tRNA are strongly protected from chemical attack. If the 3' terminus of peptidyl-tRNA binds through base pairing of

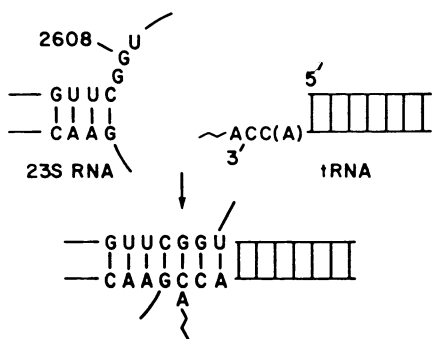


Figure 3: Model for coaxial stacking of tRNA with 23S RNA.

these two C residues, an invariant GG sequence is demanded in the ribosome. A suitable candidate is the pair of guanosines at positions 2607 and 2608. Significantly, U₂₆₀₅ is one of four possible candidates for the site of attack of the above-mentioned puromycin analogue (57). In this instance, pairing of the 3'-terminal cytosines with G₂₆₀₇ and G₂₆₀₈ creates an extended coaxial helix in which the two newly formed G-C pairs are stacked on the final C-G pair of the pre-existing helix 2588-2594/2599-2606 (Fig. 3).

(In the case of tRNA^{met}_f and the majority of other tRNA's, an additional A-U or G-U pair can be made, as shown in Fig. 3). Coaxial stacking of this kind could afford the advantage of precise structural alignment of two RNA molecules, in addition to the stability gained from the stacking itself. This type of mechanism is also attractive for the reason that it makes use of two well-known structural properties of nucleic acids: base-pairing and stacking. Such a mechanism should soon be testable.

Recently, two other models for 23S RNA secondary structure models have been proposed (67, 68). There are significant differences between these models and the one presented here. These differences will be discussed in a subsequent paper.

We wish to thank R. Hallick, I. Eperon and R. Brimacombe for sharing their unpublished results, Carol Kubitz for drawing figure 1, and M. Turner for preparing plasmid DNA. This work was supported by NIH grant GM 17129 (to H.F.N.) and NASA grant NSG 7044 (to C.R.W.).

REFERENCES

- 1 Noller, H. F. & Woese, C. R. (1981) *Science* 212, 403-411
- 2 Noller, H. F. (1980) in *Ribosomes* (eds. Chambliss, G. et al.) Univ. Park Press, Baltimore
- 3 Sanger, F. & Coulson, A. R. (1975) *J. Mol. Biol.* 94, 441-448
- 4 Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* 74, 560-564
- 5 Fox, G. & Woese, C. R. (1975) *Nature* 256, 505-506
- 6 Woese, C. R., Magrum, L. J., Gupta, R., Siegel, R. B., Stahl, D. A., Kop, J., Crawford, N., Brosius, J., Gutell, R., Hogan, J. J. & Noller, H. F. (1980) *Proc. Natl. Acad. Sci. USA* 8, 2275-2293
- 7 Brosius, J., Dull, T. J. & Noller, H. F. (1980) *Proc. Nat. Acad. Sci. USA* 77, 201-204
- 8 Kop, J., Wheaton, V., Gupta, R., Turner, M., Woese, C. R. & Noller, H. F. unpublished
- 9 Eperon, I. C., Anderson, S. and Nierlich, D. P. (1980) *Nature* 286, 460-467
- 10 VanEtten, R. A., Walberg, M. W. and Clayton, D. A. (1980) *Cell* 22, 157-170
- 11 Orozco, E. M., Rushlow, K. E., Dodd, J. R. and Hallick, R. B. (1980) *J. Biol. Chem.* 255, 10997-11003
- 12 Gourse, R. L. and Gerbi, S. A. (1980) *Nucleic Acids Res.* 8, 3623-3637
- 13 Gourse, R. L., Thurlow, D. L., Gerbi, S. A. and Zimmermann, R. A. (1981) *Proc. Natl. Acad. Sci. USA* 78, 2722-2726
- 14 Hall, L. M. C. and Maden, B. E. H. (1980) *Nucleic Acids Res* 8, 5993-6004

- 15 Sollner-Webb, B. and Reeder, R. H. (1979) *Cell* 18, 485-499
- 16 Dujon, B. (1980) *Cell* 20, 185-197
- 17 Bos, J. L., Osinga, K. A., Van der Horst, G., Hecht, N. B., Tabak, H. F., Von Ommen, G.-J. B. and Borst, P. (1980) *Cell* 20, 207-214
- 18 Allet, B. and Rochaix, J. D. (1979) *Cell* 18, 55-60
- 19 Takaiwa, F. and Sugiura, M. (1980) *Molec. gen. Genet.* 18, 55-60
- 20 Allet, B. & Spahr, P. F. (1971) *Eur. J. Biochem.* 19, 250-255
- 21 Bruce, A. G. & Uhlenbeck, O. C. (1978) *Nucleic Acids Res.* 5, 3665-3677
- 22 Contreras, R. & Fiers, W. (1979) *FEBS Lett.* 16, 281-283
- 23 Peattie, D. A. (1979) *Proc. Nat. Acad. Sci. USA* 76, 1760-1764
- 24 Herr, W. & Noller, H. F. (1978) *Biochem.* 17, 307-315
- 25 Herr, W. & Noller, H. F. (1979) *J. Mol. Biol.* 130, 421-432
- 26 Stahl, D. A. (1978) Ph.D. Thesis, Univ. of Illinois, Urbana, Illinois
- 27 Cantor, C. R. (1980) in *Ribosomes* (eds. Chambliss, G. et al.) Univ. Park Press, Baltimore, pp. 23-49
- 28 King, T. C., Rucinsky, T., Schlessinger, D. & Milanovich, F. (1981) *Nucleic Acids Res.* 9, 647-661
- 29 Edlind, T. D. & Bassel, A. R. (1980) *J. Bact.* 141, 365-373
- 30 Noller, H. F., Douthwaite, S., Garrett, R. A., & Peattie, D. A. (in preparation)
- 31 Krol, A., Machatt, M. A., Branlant, C. & Ebel, J. P. (1978) *Nucleic Acids Res.* 5, 4933-4947
- 32 Muller, R., Garrett, R. A. & Noller, H. F. (1979) *J. Biol. Chem.* 254, 3873-3878
- 33 Garrett, R. A., Douthwaite, S. & Noller, H. F. (1981) *Trends in Biochem. Sci.* 6, 137-139
- 34 Stauss, W., Gutell, R., Crawford, N. & Noller, H. F. (unpublished)
- 35 Branlant, C., Krol, A., Sriwidada, J., Sloof, P. & Garrett, R. (1976) *Eur. J. Biochem.* 70, 457-469
- 36 Branlant, C., Sriwidada, J., Krol, A. & Ebel, J. P. (1977) *Nucleic Acids Res.* 4, 4323-4345
- 37 Zimmermann, R. A. (1980) in *Ribosomes* (eds. Chambliss et al.) Univ. Park Press, Baltimore, p. 135-169
- 38 Sloof, P., Hunter, J. B., Garrett, R. A. & Branlant, C. (1978) *Nucleic Acids Res.* 5, 3503-3513
- 39 Maly, P., Rinke, J., Ulmer, E., Zwieb, C. & Brimacombe, R. (1980) *Biochemistry* 19, 4179-4188
- 40 Herr, W. & Noller, H. F. (1975) *FEBS Lett.* 53, 248-252
- 41 Branlant, C., Krol, A., Sriwidada, J. & Brimacombe, R. (1976) *Eur. J. Biochem.* 70, 483-492
- 42 Sloof, P. (1977) Ph.D. Thesis, Univ. of Amsterdam
- 43 Wrede, P. & Erdmann, V. A. (1977) *Proc. Nat. Acad. Sci. USA* 74, 2706-2709
- 44 Nazar, R. N. (1980) *FEBS Lett.* 119, 212-214
- 45 Luehrsen, K. R. & Fox, G. E. (1981) *Proc. Nat. Acad. Sci. USA* 78, 2150-2154
- 46 Jordan, B. R. (1974) *FEBS Lett.* 44, 39-42
- 47 Pace, N. R., Walker, T. A. & Schroeder, E. (1977) *Biochem.* 16, 5321-5328
- 48 Pavlakis, G. N., Jordan, B. R., Wurst, R. M. & Vournakis, J. N. (1979) *Nucleic Acids Res.* 7, 2213-2238
- 49 MacKay, R. M. (1981) *FEBS Lett.* 123, 17-18
- 50 Nomiya, H., Sakaki, Y. and Takagi, Y. (1981) *Proc. Natl. Acad. Sci. USA* 78, 1376-1380
- 51 Rae, P. M. M., Kohorn, B. D. & Wade, R. P. (1980) *Nucleic Acids Res.* 8, 3491-3504
- 52 Roiha, H., Miller, J. R., Woods, L. C. & Glover, D. M. (1981) *Nature* 290, 749-753
- 53 Wild, M. A. & Sommer, R. (1980) *Nature* 283, 693-694

-
- 54 Gilbert, W. (1978) *Nature* 271, 501
- 55 Cooperman, B. S. (1980) in *Ribosomes*, eds. Chambliss et al., Univ. Park Press, Baltimore, pp. 531-554
- 56 Yukioka, M., Hatayama, T. & Omari, K. (1977) *Eur. J. Biochem.* 73, 449-459
- 57 Eckermann, D. J. & Symons, R. H. (1978) *Eur. J. Biochem.* 82, 225-234
- 58 Cundliffe, E. (1980) in *Ribosomes*, eds. Chambliss, G. et al., Univ. Park Press, Baltimore, pp. 555-581
- 59 Kearsay, S. E. & Craig, I. W. (1981) *Nature* 290, 607-608
- 60 Monro, R. E., Cerná, J. & Marcker, K. A. (1968) *Proc. Nat. Acad. Sci. USA* 61,
- 61 Cerná, J., Rychlik, I., Krayevsky, A. A. and Gottikh, B. P. (1974) *Acta Biol. Med. Ger.* 33, 877.
- 62 Peattie, D. A. & Herr, W. (1981) *Proc. Nat. Acad. Sci. USA* 78, 2273-2277
- 63 Spierer, P., Zimmermann, R. A. & Branlant, C. (1976) *FEBS Lett.* 68, 71-75
- 64 Selker, E. and Yanofsky, C. (1979) *Nucleic Acids Res.* 6, 2561-2567
- 65 Fanning, T. G. and Traut, R. R. (1981) *Nucleic Acids Res.* 9, 993-1004
- 66 Edwards, K. & Körsel, H. (1981) *Nucleic Acids Res.* 9, 2853-2869
- 67 Glotz, C., Zwieb, C. & Brimacombe, R. (1981) *Nucleic Acids Res.* 9, 3287-3306
- 68 Branlant, C., Krol, A., Machatt, M. A., Pouyet, J. & Ebel, J. B. (1981) *Nucleic Acids Res.*, in press.