Volume 11 Number 22 1983 Nucleic Acids Research

Secondary structure of the *Dictyostelium discoideum* small subunit ribosomal RNA*

Gary J.Olsen, Robert McCarroll and Mitchell L.Sogin

Department of Molecular and Cellular Biology, National Jewish Hospital and Research Center, 3800 East Colfax Avenue, and Department of Biochemistry, Biophysics and Genetics, University of Colorado Health Sciences Center, Denver, CO 80206, USA

Received 29 August 1983; Revised and Accepted 17 October 1983

ABSTRACT

We have used comparative analyses of prokaryotic and eukaryotic small subunit ribosomal RNAs to deduce a secondary structure for the Dictyostelium discoideum 18S rRNA. Most of the duplex regions are evolutionarily conserved in all organisms. We have taken advantage of the variation to the D. discoideum sequence (relative to the yeast and frog 18S rRNAs) to identify additional helical regions which are common to the eukaryotic 18S rRNAs.

INTRODUCTION

Knowledge of ribosomal RNA (rRNA) secondary structures will broaden our understanding of protein synthesis. Given a single sequence, free energy rules (1) and/or empirical rules (2) can be used to estimate the most favorable structure. In recent years considerable progress has been made toward the prediction and confirmation of the higher-order structures of the 5S and 5.8S rRNA sequences, however similar analyses for the larger nucleic acid components of the ribosome, i.e. the 16-18S and 23-28S rRNAs, are less advanced. Computer searches of the Escherichia coli small subunit rRNA reveal 10,000 possible helices of four or more base pairs (3); fewer than 100 of which can simultaneously exist. There is little hope of ever refining a set of free energy estimates to sufficiently resolve those numerous possibilities.

An alternative approach when several functionally homologous sequences are known is identification of phylogenetically conserved secondary and tertiary structural features. Such analyses have been used to infer transfer RNA (4,5), 5S rRNA (6,7,8), and 5.8S rRNA (8,9,10) foldings. The transfer RNA structure has been confirmed by X-ray crystallography (11), while the 5S and 5.8S rRNA foldings are supported by chemical modification and nuclease sensitivity mapping experiments (12,13,14). A picture of the small subunit rRNA (generic 16S-18S rRNA) secondary structure is also emerging. Noller and Woese (3,15) have presented a consensus folding for the prokaryotic 16S rRNA which is consistent with the eubacterial, archaebacterial, and organellar 16S rRNA sequences and the T_1 RNase oligonucleotide catalogues from over 150

prokaryotes. In addition, Stiegler et al. (16) have identified sequence complementaries which are common to the prokaryotic as well as two eukaryotic (yeast and frog) 18S rRNA sequences; however the limited data did not allow the identification of presumptive helices which are unique to the eukaryotes.

Recently we have sequenced the entire Dictyostelium discoideum 18S rRNA gene (17). The inferred RNA sequence represents the deepest divergence in the eukaryotic line of descent yet characterized by molecular phylogeny (10,17). Taking advantage of this sequence divergence, it has been possible to identify additional sequence complementarities which are common to the eukaryotic 18S rRNAs. Here we present a preliminary secondary structure for the D. discoideum 18S rRNA which is consistent with possible foldings of the corresponding Saccharomyces cerevisiae and Xenopus laevis sequences. Most of the proposed duplex regions are evolutionarily conserved in all organisms, however a few can be constructed only in the eukaryotic 18S rRNA sequences. We have identified those helices which can be formed despite variations in the primary structure and thusly are considered to be phylogenetically proven. Additional sequences from distantly related eukaryotes will be required in order to confirm several unproven eukaryote-specific structures proposed in the model.

DERIVATION OF THE MODEL

The D. discoideum small subunit rRNA secondary structure model was inferred through a comparative analysis of sequences. The fundamental hypothesis of the analysis is "functionally important sequence features are conserved through evolution". Using this approach Stiegler et al. (16) have identified many potential pairings common to the eubacterial and the S. cerevisiae and X. laevis small subunit rRNA sequences. Because of the tremendous phylogenetic separation of the kingdoms and the extra length of nuclear defined eukaryotic 18S rRNAs relative to their bacterial and organellar counterparts, we anticipated the existence of additional "eukaryote-specific" secondary structure.

The identification of helices as homologous or nonhomologous is dependent upon the choice of sequence alignment. The small subunit rRNA gene sequences from D. discoideum (17), X. laevis (18), and S. cerevisiae (19) nuclei, as well as those from E. coli (20), Proteus vulgaris (a direct RNA sequence) (21), Halobacterium volcanii (22), three chloroplasts (23,24,25), and several mitochondria (26,27,28,29,30) were initially aligned with one another on the basis of primary structural homologies. The relative alignments of the S. cerevisiae, X. laevis and D. discoideum sequences were then refined on the

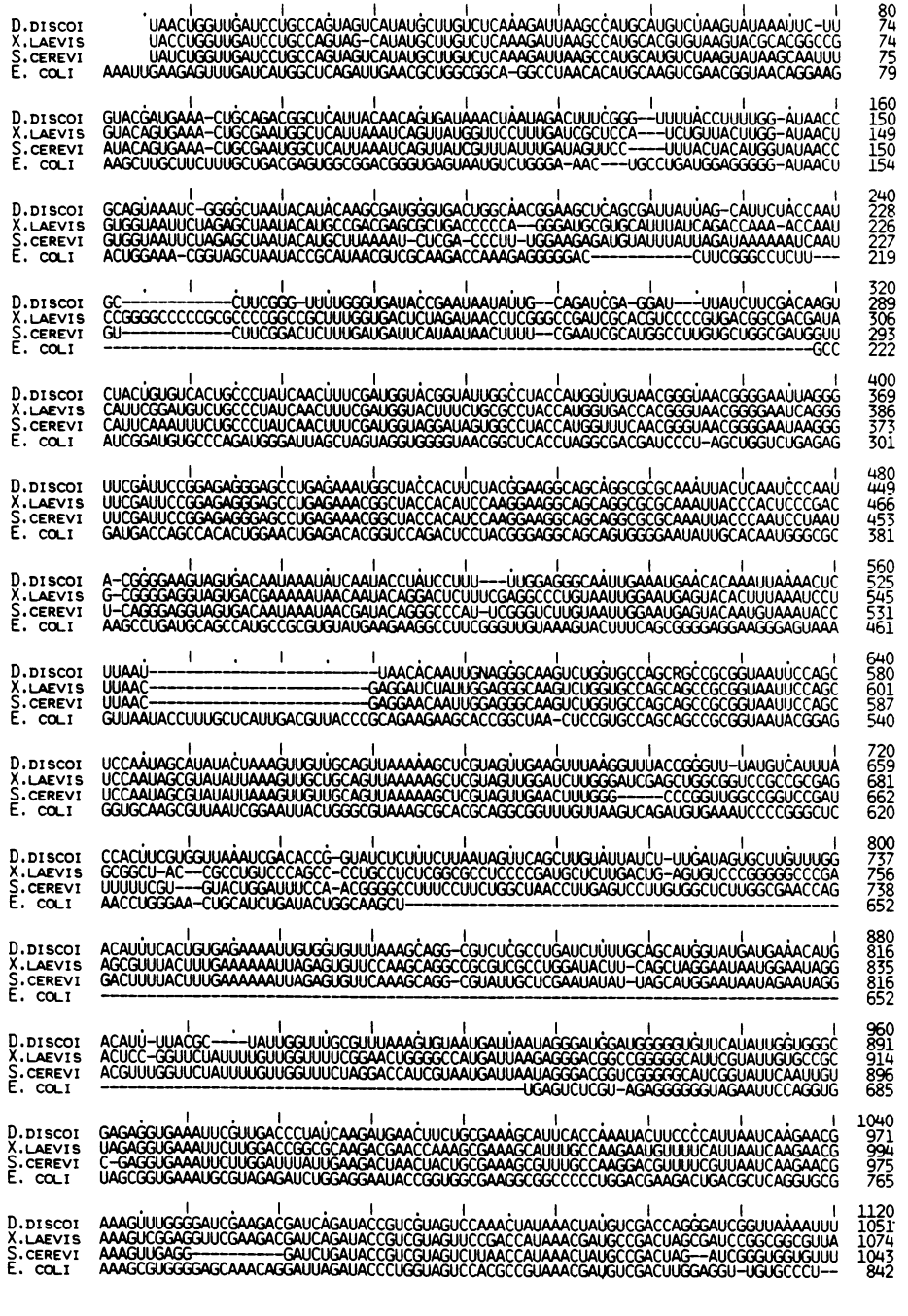
basis of sequence homologies which are unique to the eukaryotes (17). This process was repeated as additional but less extensive homologies were located. Finally, as conserved secondary structural features were identified, they provided additional landmarks in regions of otherwise ambiguous alignment. For example, there are several instances in which regions of sequence length variation could be localized to the loops of "hairpins" in the consensus folding; the rRNAs from the various sources may have very divergent primary structures and different loop lengths, yet maintain homologous stem locations. Similarly, if one half of a conserved duplex could be unambiguously aligned among the rRNAs, then the sequences defining the second half of the duplex could be aligned on the basis of the pairing. Thus, the alignment was continuously refined as the analysis progressed. Figure 1 presents the final alignment of the D. discoideum, X. laevis, S. cerevisiae and E. coli sequences.

We started constructing the D. discoideum folding by assuming that the secondary structures for the eukaryotic and prokaryotic rRNAs are similar (15,16,31). Therefore we initially examined the eukaryotic sequences for potential pairings which are analogous to those found in the eubacterial model proposed by Noller and Woese (3).

A computer program (G.J.O., unpublished) was used to scan the remaining unpaired regions for additional complementary sequences in the eukaryotic 18S rRNAs. The presumptive helices fall into two categories; those which are defined by regions of sequence variation, and thus are phylogenetically proven (see below), or those which are defined by regions of little or no sequence variation, and thus lack proof of secondary structure. A helical region is considered to be proven if its formation is independent of primary structure; compensating base changes must be found which maintain sequence complimentarity. We considered three sources of sequence variation when evaluating the phylogenetic evidence for a given helix: compensated variation within the eukaryotes (eukaryotic proof), variation within the eubacteria (eubacterial proof), and variation between kingdoms (interkingdom proof) (15). Interkingdom evidence frequently is a redundant measure when proof exists within the eukaryotes or the eubacteria, however in some cases the degree of interkingdom proof is greater than the evidence within a kingdom.

RESULTS AND DISCUSSION

Table I lists the locations and summarizes the phylogenetic evidence for duplex regions in our D. discoideum 18S rRNA secondary structure model. We



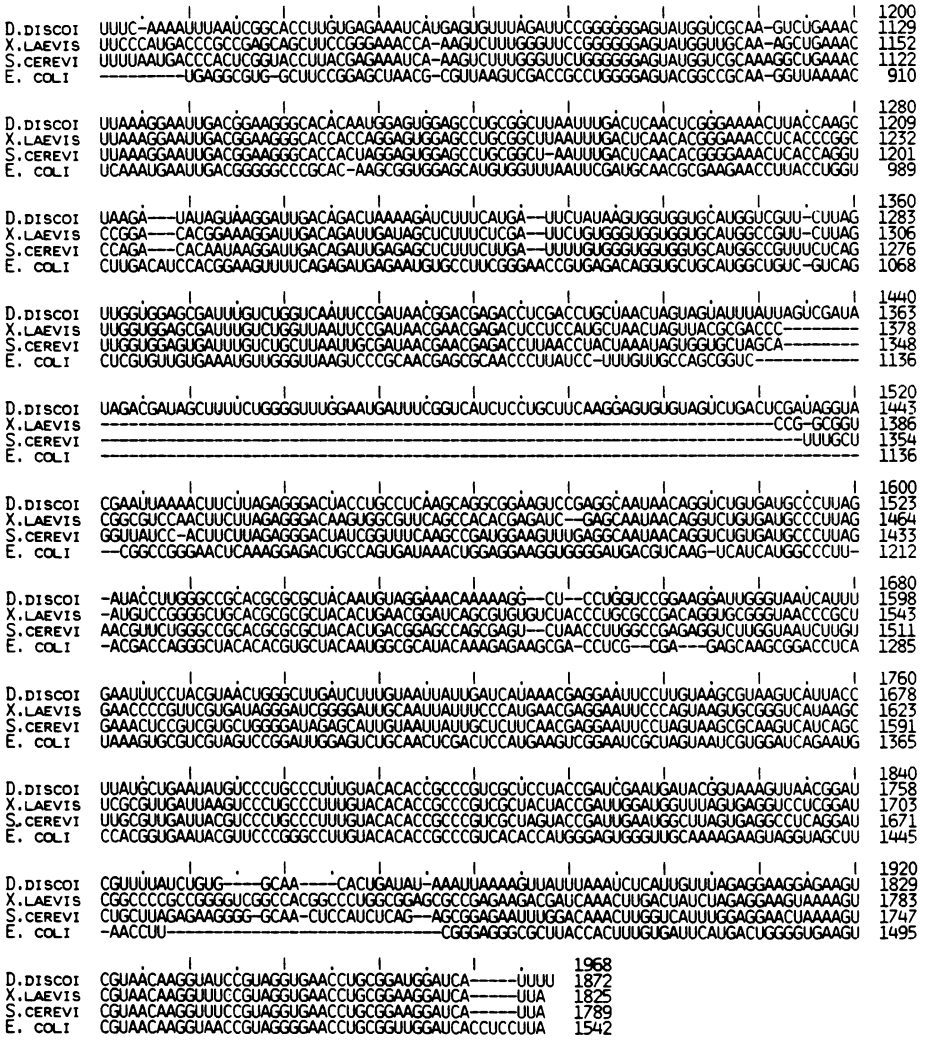


Figure 1. The Sequence of the Dictyostelium discoideum Small Subunit Ribosomal RNA Coding Region Aligned with Other Small Subunit rRNAs. The sequence of the D. discoideum small subunit rRNA (17) is shown aligned with those from Xenopus laevis (18), Saccharomyces cerevisiae (19), and Escherichia coli (20). Initially the sequences were aligned according to primary structure. The locations of evolutionarily conserved secondary structures were then used to refine the alignment where length variation occurred. The differences in sequence lengths were compensated by introducing appropriate gaps (-) into the sequences. Nucleotide numbering for each sequence is provided at the right margin. To facilitate locating the helical regions in Table I, a uniform numbering (corresponding to the "aligned positions" in the third column of the table) is included above the D. discoideum sequence.

have indicated in Table I whether pairings homologous to those in our D. discoideum model can be accommodated by other eukaryotic sequences or by the eubacterial sequences (including the organellar sequences). The two dimensional folding is shown in Figure 2. The major structural regions of the model are similar to those found in the eubacterial folding proposed by Noller and Woese (3,15) and the S. cerevisiae 18S rRNA model proposed by Mankin, et al. (31). These regions are referred to as the 5', the middle, and the 3' domains, corresponding to positions 1-600, 601-1140, and 1141-1872, respectively. We wish to call attention to interesting features within the structure. These include helices which are present in the eukaryotes but are not found in the eubacterial models, helices for which the variation supplied by D. discoideum sequence was either essential or contributed strongly to the structural proof, and helices which are not energetically favored but are phylogenetically proven in the 18S rRNA consensus folding.

The 5' domain is a composite of universal helices (structures which are found in both the eubacterial and eukaryotic foldings) and eukaryote-specific duplex regions (pairings which cannot be accommodated by the bacterial sequences). Most of the helices between positions 112-297 are phylogenetically proven. Helix 8 (140-154/159-174) is universal and contains a number of unusual base pairs interspersed with the positions of proven pairing. Helix 9 (178-182/259-263), helix 10 (183-191/196-204), and helix 13 (267-273/278-284) are well-proven, eukaryote-specific structures; the latter two isolate regions of length variation in their hairpin loops. The sequence variation of the D. discoideum 18S rRNA provides the evidence for helix 13. This region contains two other eukaryote-specific helices, 11 (209-214/253-258) and 12 (223-227/241-245), which are not as well-proven. The lack of proof for helix 12 is a reflection of an ambiguity in the alignment of this portion of the X. laevis sequence.

This region of the D. discoideum sequence can also pair UAGACUU (120-126) with AAGUCUA (286-292). The 16S rRNAs from both kingdoms of the prokaryotes form an analogous, well-proven helix (15). However, because the S. cerevisiae and X. laevis sequences cannot accommodate this pairing, we have not displayed the helix in our model. This structure may, in fact, be an instance where D. discoideum forms a "prokaryote-specific" pairing which is absent in other eukaryotic foldings.

A second region (positions 467-536) in the 5' domain lacks primary and secondary structural homologies with the eubacterial sequences. With the exception of the initial 4 basepairs, helix 20 (474-486/492-504) is well

Table I: Phylogenetic evidence for helical regions in the Dictyostelium discoideum small subunit ribosomal RNA secondary structure model [a]

helical region	position		presence		proof			
	D.d. [b]	aligned [c]	Euk [d]	Eub [e]	Euk [f]	Eub [g]	King [h]	D.d. sequence [i]
1	4-8 16-20	(9-13) (21-25)	+	+	-	+++		CUGGU GACCG
2	12-15 1134-1137	(17-20) (1204-1207)	+	+	-	+++		UCCU AGGA
3	21-32 587-597	(26-37) (646-656)	+	+	-	+++		UAGUcAUAUGCU AUCA-UAUACGA
4	33-36 463-466	(38-41) (494-497)	+	-	-			UGUC ACAG
5	48-54 422-427	(53-59) (452-457)	+	+	-	-	+	GCCaUGC CGG-ACG
6	63-68 73-78	(68-73) (79-84)	+	-	+/-			GUAUAA CAUGUU
7	106-112 297-303	(113-119) (327-333)	+	-	+/-			CAGUGAU GUCACUG
8	140-154 159-174	(149-164) (169-185)	+	+	+	+++		UuUGgA-UAaCCgCaG AuACaUaAUcGGgGcU
9	178-182 259-263	(189-193) (283-289)	+	-	+++			GCGAU CGUUA
10	183-191 196-204	(194-202) (207-215)	+	-	+++			GGGUgaCUG CUCGaaGGC
11	209-214 253-258	(220-225) (277-282)	+	-	+/-			AUUAUU UAAUAA
12	223-227 241-245	(235-239) (265-269)	+	+/-	+/-			ACCaA UGGgU
13	267-273 278-284	(293-300) (308-314)	+	-	++			UCGAGGA AGCUUCU
14	304-306 351-353	(334-336) (381-383)	+	+	-	?	+	CCC GGG
15	319-328 333-341	(349-358) (363-371)	+	+	-	+++		AUGGUAcGGU UACCAU-CCG
16	357-364 369-379	(387-394) (399-409)	+	+	-	+++		CGGGG---AAU GCCUUagcUUG
17	384-390 399-405	(414-420) (429-435)	+	+	-	+++		GGgAGCC CCaUCGG
18	407-410 415-418	(437-440) (445-448)	-	+		+++		CUUC GAAG
19	434-446 452-462	(464-476) (483-493)	+	+	+	+++		AUUACUcaaUCCC UGAUGAa--GGGG
20	474-486 492-504	(505-517) (526-538)	+	+	+/-	+		UCAA-UaCCUaUCC AGUUAacGGG-AGG
21	515-520 526-531	(549-554) (560-590)	-	-				AAUUAA UUAAUU
22	538-544 580-586	(597-603) (639-645)	+	+	-	+++		AUUGGAG UAACCUC
23	552-557 574-579	(611-616) (633-638)	+	+	-	+++		CUGGuG GACCuU
24	600-603 1029-1032	(659-662) (1097-1100)	+	-	-			GUUG CAGC
25	619-622 969-972	(678-681) (1037-1040)	+	+/-	-			UCGU AGCA
26	817-827 839-849	(880-891) (907-917)	+	-	+++			ACAUUUUAcGC UGUGAAAAUuUG
27a	852-859 958-965	(920-927) (1026-1033)	+	-	-			UGAUUAAU ACUAAUUA
27b	861-876 942-956	(929-944) (1010-1024)	+	+	++	+++		GGGAuGgAUgggGGUG CCCUuCaUAaa-CCAC
28a	877-881 921-925	(945-949) (989-993)	+	-	++			UUCAU AAGUA
28b	883-890 912-919	(951-958) (980-987)	+	+	+++	+++		UUGGUGGG AACUAUCC
29	894-896 903-905	(962-964) (971-973	+	+	-	+++		GAG CUU

Table I (continued)

helical region	position D.d.[b]	position aligned[c]	presence Euk[d]	presence Eub[e]	proof Euk[f]	proof Eub[g]	proof King[h]	D.d. sequence[i]
30	974-981	(1042-1049)	+	+	+++	+++		AGUUUGGG
	1010-1017	(1078-1085)						UCAAACCU
31	989-992	(1057-1060)	+	+	-	+++		GACG
	1002-1005	(1070-1073)						CUGC
32	1034-1052	(1102-1120)	+	+	+++	+++		AGGgaUCGGUUAAaAUUUU
	1056-1073	(1125-1142)						UCCacGGCUAAUU-UAAAA
33	1081-1088	(1150-1157)	+	-	+/-			AAUCaUGA
	1093-1099	(1162-1168)						UUAG-AUU
34	1107-1116	(1176-1185)	+	+	+/-	+++		GAGUa-UGGuC
	1121-1131	(1191-1201)						UUCAaaGUCuG
35	1142-1152	(1212-1222)	+	+	-	+++		ACgGAAGGGCA
	1697-1706	(1778-1787)						UG-UUUCCCGU
36	1159-1163	(1229-1233)	+	+	-	+++		GGAGU
	1653-1657	(1734-1738)						CCUUA
37	1166-1175	(1236-1245)	+	+	-	+++		AGCcUGCG-GC
	1536-1546	(1613-1623)						UCGcGCGCaCG
38	1180-1184	(1250-1254)	+	+	-	+/-	+	UUUGA
	1191-1195	(1261-1265)						GGGCU
39	1204-1211	(1274-1281)	+	+	++	+++		CCAAGcUA
	1525-1532	(1602-1609)						GGUUCcAU
40	1215-1219	(1288-1292)	+	+	++	+++		UAUAG
	1254-1258	(1329-1333)						AUAUC
41	1227-1235	(1300-1308)	+	+/-	-	-	+/-	UGAcAGA-CU
	1239-1248	(1312-1321)						ACUuUCUaGA
42	1262-1278	(1337-1354)	+	+	-	++		GGuGGUG-CAUGG-UC-GUU
	1502-1520	(1578-1596)						UC-CCGUaGUGUCuGGaCAA
43	1283-1285	(1359-1361)	+	+	-	+/-	+/-	GUU
	1320-1322	(1396-1398						CAG
44	1286-1291	(1362-1367)	+	+	-	++		GGUGGA
	1296-1301	(1372-1377)						CUGUUU
45	1302-1304	(1378-1380)	+	+	+/-	++		UGG
	1311-1313	(1387-1389)						GCC
46	1327-1331	(1403-1407)	+	+/-	++	-		CCUCG
	1493-1497	(1569-1573)						GGAGC
47	1333-1340	(1409-1416)	+	+	-	+++		CCUgCUAA
	1458-1464	(1534-1540)						GGA-GAUU
48	1347-1395	(1423-1471)	-	-				D.d. specific insert
	1400-1444	(1476-1520)						
49	1470-1474	(1546-1550)	+	+	+++	+		CCUGC
	1480-1484	(1556-1560)						GGACG
50	1550-1559	(1627-1636)	+	+	+++	+++		AUGUAGGAAA
	1602-1611	(1683-1692)						UGCAUCCUUU
51	1570-1577	(1651-1658)	+	-	+			CCUGGUCC
	1582-1589	(1663-1670)						GGGUUAGG
52	1621-1625	(1702-1706)	+	+	++	+++		UGAUC
	1638-1642	(1719-1723)						ACUAG
53	1663-1670	(1744-1751)	+	+	+	+++		AGCGUAAG
	1678-1685	(1759-1766)						UCGUAUUC
54	1722-1771	(1803-1852)	+	+	+++	+++		CUCCUaCCgaUcGAAUGAU...
	1776-1825	(1866-1916)						GAGGAaGGa-GaUUUGUUA...
55	1840-1849	(1931-1940)	+	+	+	+++		UAUCCGUAGG
	1854-1863	(1945-1954)						GUAGGCGUCC

[a] The aligned small subunit rRNA sequences (see text) were used to identify evolutionarily conserved helical regions. A semi-quantitative measure for phylogenetic proof (see text) is provided: "-" is unproven; "+/-" means very limited proof (one example of sequence variation); "+" is partial proof (two compensated changes with no counter examples); "++" is good proof (multiple examples of compensated sequence variation); and "+++" corresponds to very good proof (numerous examples of compensated sequence variation).

[b] The endpoints of the paired regions in the D. discoideum small subunit rRNA sequence. This is the numbering system of Figure 2.

[c] The endpoints of the paired regions in the aligned sequence numbering system of Figure 1.

[d] Helical regions which are present in eukaryotic consensus foldings. Helices indicated as being absent (-) in the consensus folding are unique to the D. discoideum secondary structure model.

[e] Helical regions which are also present in the eubacterial consensus foldings. Helices which cannot be unequivocally identified in the eubacterial foldings are indicated with "+/-".

[f] Extent of eukaryotic proof (see text).

[g] Extent of eubacterial proof (see text).

[h] Extent of interkingdom proof (see text).

[i] The nucleotide sequence of the pairing region in the D. discoideum small subunit rRNA sequence. The top line of sequence reads from 5' to 3' in the sequence, the bottom line is reversed. Orthodox base pairs are shown in upper-case; bulges and mismatched pairs are lower case.

proven within the eukaryotes. In contrast helix 21 (515-520/526-531) can only be formed in the D. discoideum 18S rRNA; S. cerevisiae and X. laevis can form shorter helices in the region but their alignment with the D. discoideum helix is not precise.

The central region of the 5' domain can assume one of two foldings. An alternative to the displayed structure extends helix 14 (304-306/351-353) with UG/UA (302-303/354-355). This disrupts helix 7 (106-115/297-303), but permits the formation of a new eukaryote-specific helix, ACUG (85-88) paired with CAGU (106-109). Neither of these alternatives can be proven with the available data.

There are three other helices in the 5' domain which we wish to discuss. Helix 6 (63-68/73-78), a partially proven eukaryote-specific helix, defines a region of length variation in its hairpin loop. Helix 18 (407-410/415-418) is well proven in the eubacteria, but the X. laevis and S. cerevisiae 18S rRNAs form A/A mismatches within the helix. This suggests that this region of the D. discoideum 18S rRNA is not typically eukaryotic, but contains some features of the eubacterial structure. Helix 16 (357-364/369-379) is a well-proven, universal helix which contains a three nucleotide (CGA in the eukaryotes) bulge. Although the existence of the bulge in the eubacterial sequences contradicts the structure predicted from free energy rules (1,2,32), the pairing presented is supported by at least five perfectly compensated sequence variations.

The middle domain displays a remarkable range of evolutionary constraints. This is evidenced by a lengthy region of nonconserved primary and secondary structure followed by a region of extreme conservation. Consequently, some structures which may be functionally equivalent are difficult to identify and align. For example the unproven helix 25 (619-622/969-972) leads into a region (positions 626-810) which displays extreme sequence variation in all small subunit rRNAs. The eubacteria have a pairing which may

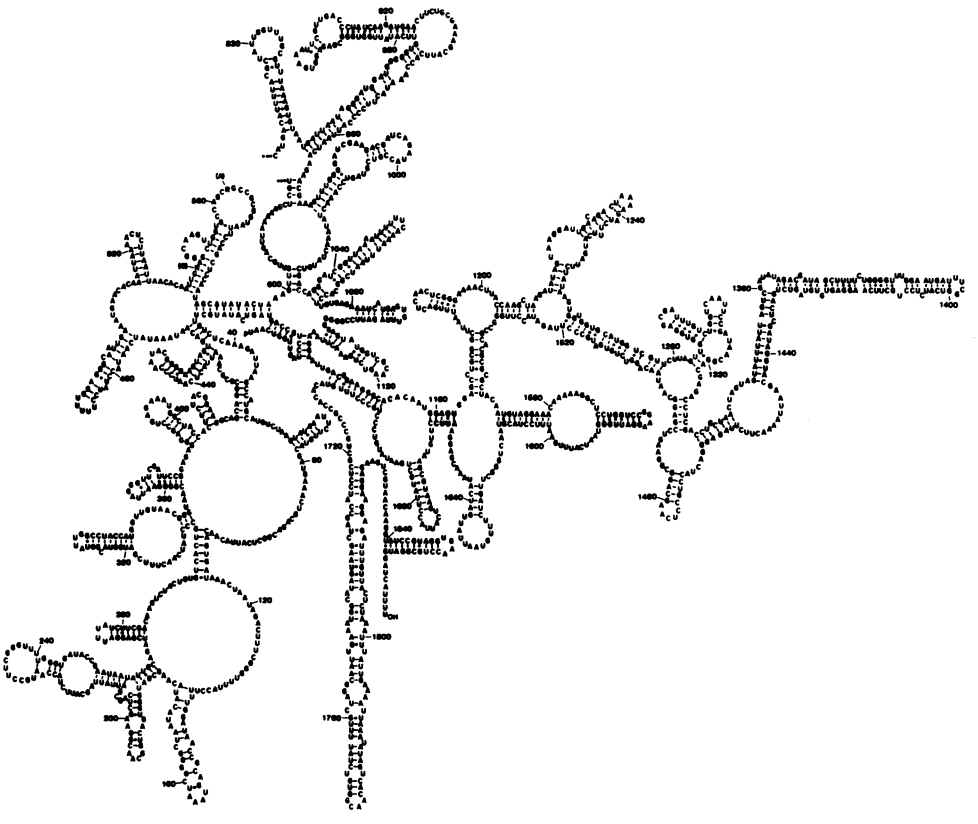


Figure 2. Secondary Structure of the Dictyostelium discoideum Small Subunit rRNA. The secondary structure is based upon the duplex regions which are listed in Table I.

be equivalent to helix 25. The adjacent variable region is the binding site for the S8 ribosomal protein (33,34). This region in the eubacterial sequences can be folded into a long, unbranched stem (with appropriate internal loops and bulges (15)). In our model the corresponding region is not shown because there is no consensus folding for the available eukaryotic sequences. Rather than speculate on the structure of this portion of the molecule we prefer to wait until additional sequence data become available.

At the termination of the S8 region is a eukaryote-specific helix, 26 (817-827/839-849), whose existence is strongly supported by variation in the D. discoideum sequence. Adjacent to helix 26 is the 27a/27b helical region. Helix 27a (852-859/958-965) is an unproven, eukaryote-specific extension of

the universal helix 27b (861-876/942-956). Helix 27b is well-proven in the eukaryotes by variation in the D. discoideum sequence. Similarly, helix 28a (877-881/921-925) is an unproven eukaryote-specific extension of the universal helix 28b (883-890/912-919). Helix 28b leads into a new universal helix, 29 (894-896/903-905), which is supported by the mitochondrial small subunit rRNA sequence diversity and interkingdom sequence variations.

The final noteworthy feature of the middle domain is an ambiguity associated with helices 24 (600-603/1029-1032) and 33 (1081-1088/1093-1099). An alternative structure is the formation of a eukaryote-specific helix, GAU (1096-1098) paired with GUC (1027-1029), disrupting helices 24 and 33. Both alternatives are supported by a single compensated base change. We have displayed helices 24 and 33 in our model only because they result in a greater total number of basepairs. The alternative to the displayed pairings is similar to, but much shorter than, a proven eubacterial pairing. Additional data will be required to resolve the issue.

In general the 3′ domain is a collection of universal structures, some of which display minor, kingdom-specific variations. Helix 39 (1204-1211/1525-1532) is a universal pairing which contains a eukaryote-specific prymidine/prymidine mismatch. The eukaryotic structural proofs for this helix and for the adjacent universal helix, 40 (1215-1219/1254-1258), are provided by the D. discoideum sequence. Helices 43 (1283-1285/1320-1322) and 44 (1286-1291/1296-1301) can be considered as a nine base pair universal structure, however the transition from helix 43 to helix 44 occurs at a "kingdom-specific" location.

There are a few pairings in the 3′ domain which appear to be present only in the eukaryotic folding. Helix 41 (1227-1235/1239-1248) is an unproven eukaryote-specific pairing. The analogous region in H. volcanii can also pair, but ambiguity in the sequence alignment makes it difficult to evaluate the significance of this observation. Helices 46 (1327-1331/1493-1497) and 51 (1570-1577/1582-1589) are well-proven eukaryote-specific pairings. Some of the eubacterial sequences can pair in similar regions, but they do not display compensating base changes when one of the "pairing partners" changes.

A feature in the 3′ domain which is unique to the D. discoideum sequence is helix 48 (1347-1395/1400-1444). It represents an insert of approximately 80 nucleotides relative to the other eukaryotic sequences. Because this region can pair a remarkably large fraction of its nucleotides (44 base pairs compared with 35-37 in the D. discoideum 5S rRNA (35) which is almost identical in length) we have displayed the unproven structure.

Finally we wish to call attention to the penultimate helix, 54 (1722-1771/1776-1825). This structure is well-proven in all kingdoms, but it is difficult to draw a universal folding. In part, this is due to the numerous bumps and bulges which must be included in the pairing. Among the eukaryotic sequences there are proven pairings distributed along the entire length of the arm.

Throughout our analysis of D. discoideum 18S rRNA secondary structure, we noted the general conservation of primary structure among the eukaryotic 18S rRNAs sequences. Of the 57 helical regions in the model, only 29 could be supported by variation among the eukaryotic sequences (see Table I). The importance of additional sequence data is emphasized by noting that the inclusion of the D. discoideum sequence in the comparisons provided the eukaryotic proof for six (21%) of the 29 pairings, and it strengthened the support for an additional 17 helices.

The remaining helices in the model are either unproven or relied upon variation in the other kingdoms for their support. In Figure 1, 747 nucleotides lie in regions of five or more consecutive positions which lack eukaryotic sequence variation; consequently, secondary structures within these regions cannot be proven within the eukaryotic kingdom. If these regions in the D. discoideum sequence could accommodate the eubacterial pairings, we chose to accept them on the basis of interkingdom structural homology. If the eukaryotic sequences could not be accurately fit to the eubacterial pairings, then the choice of structure must be considered speculative. Resolution of these speculations will require additional sequence data from phylogenetically diverse eukaryotes.

ACKNOWLEDGEMENTS

The authors thank Drs. R. Gupta and C.R. Woese for providing us with the Halobacterium volcanii small subunit rRNA sequence prior to publication. We thank Drs. C.R. Woese and Norman Pace for critical discussion of the proposed Dictyostelium discoideum small subunit rRNA folding.

*This investigation was supported by National Institutes of Health, Research Grant GM23464 to M.L.Sogin

BIBLIOGRAPHY

1. Tinoco, I., Jr., Borer, P.N., Dengler, B., Levine, M.D., Uhlenbeck, O.C., Crothers, D.M. and Gralla, J. (1973). Nature New Biol. 246:40-41.
2. Ninio, J. (1979). Biochimie 61:1133-1150.

3. Noller, H.F. and Woese, C.R. (1981). Science 212:403-411.
4. Gauss, D.H. and Sprinzl, M. (1983). Nucl. Acids Res. 11:r1-53.
5. Rich, A. and RajBhandary, U.L. (1976). Ann. Rev. Biochem. 45:805-860.
6. Fox, G.E. and Woese, C.R. (1975). Nature (London) 256:505-507.
7. Hori, H. and Osawa, S. (1979). Proc. Natl. Acad. Sci. USA 76:381-385.
8. MacKay, R.M., Spencer, D.F., Schnare, M.N., Doolittle, W.F. and Gray, M.W. (1982). Can. J. Biochem. 60:480-489.
9. Ursi, D., Vandenberghe, A. and De Wachter, R. (1982). Nucl. Acids Res. 10:3517-3530.
10. Olsen, G.J. and Sogin, M.L. (1982). Biochemistry 21:2335-2343.
11. Kim, S.-H. (1979). in Transfer RNA: Structure, Properties and Recognition Schimmel, P.R., Söll, D., and Abelson, J.N., eds. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory), pp83-100.
12. Walker, T.A., Johnson, K.D., Olsen, G.J., Peters, M.A. and Pace, N.R. (1982). Biochemistry 21:2320-2329.
13. Nazar, R.N., Sitz, T.O. and Busch, H. (1975). J. Biol. Chem. 250:8591-8597.
14. Kelly, J.M. and Maden, B.E.H. (1980). Nucl. Acids Res. 8:4521-4534.
15. Woese, C.R. and Noller, H.F. (1983). Microbiol. Rev. in press.
16. Stiegler, P., Carbon, P., Ebel, J.-P. and Ehresmann, C. (1981). Eur. J. Biochem. 120:487-495.
17. McCarroll, R., Olsen, G.J., Stahl, Y.B., Woese, C.R. and Sogin, M.L. (1983). The Nucleotide Sequence of the Dictyostelium discoideum Small Subunit Ribosomal RNA Inferred from the Gene Sequence: Evolutionary implications. Biochemistry, in press.
18. Salim, M. and Maden, B.E.H. (1981). Nature (London) 291:205-208.
19. Rubstov, P.M., Musakhanov, M.M., Zakharyev, V.M., Krayev, A.S., Skryabin, K.G. and Bayev, A.A. (1980). Nucl. Acids Res. 8:5779-5794.
20. Brosius, J., Palmer, M.L., Kennedy, P.J. and Noller, H.F. (1978). Proc. Natl. Acad. Sci. USA 75:4801-4805.
21. Carbon, P., Ebel, J.P. and Ehresmann, C. (1981). Nucl. Acids Res. 9:2325-2333.
22. Gupta, R., Lanter, J.M. and Woese, C.R. (1983). Science. 221:656-659.
23. Graf, L., Roux, E., Stutz, E. and Kössel, H. (1982). Nucl. Acid Res. 10:6369-6381.
24. Schwarz, Zs. and Kössel, H. (1980). Nature (London) 283:739-742.
25. Dron, M., Rahire, M. and Rochaix, J.-D. (1982). Nucl. Acids Res. 10:7609-7619.
26. Küntzel, H. and Köchel, H.G. (1981). Nature (London) 293:751-755.
27. Sor, R. and Fukuhara, H. (1980). C.R. Acad. Sci. (Paris) D291:933-936.
28. Seilhamer, J.J., Olsen, G.J. and Cummings, D.J. (1983). Paramecium Mitochondrial Genes: 1. Small subunit rRNA Gene Sequence and Microevolution. Submitted for publication.
29. Eperon, I.C., Anderson, S. and Nierlich, D.P. (1980). Nature 286:460-467.
30. Van Etten, R.A., Walberg, M.W. and Clayton, D.A. (1980). Cell 22:157-170.
31. Mankin, A.S., Kopylov, A.M. and Bogdanov, A.A. (1981). FEBS Letters 134:11-14.
32. Salser, W. (1977). Cold Spring Harbor Symp. Quant. Biol. 42:985-1002.
33. Zimmerman, R.A. and Singh-Bergmann, K. (1979). Biochim. Biophys. Acta 563:422-431.
34. Müller, R., Garrett, R.A. and Noller, H.F. (1979). J. Biol. Chem. 254:3873-3878.
35. Hori, H., Osawa, S. and Iwabuchi, M. (1980). Nucl. Acids Res. 8:5535-5539.