Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 5044–5048, June 1992
Biochemistry

Higher order interactions in 23S rRNA

NIELS LARSEN*

Chemistry Department, Aarhus University, Langelandsgade 140, 8000 Aarhus C, Denmark

Communicated by Carl R. Woese, January 3, 1992

ABSTRACT **An alignment of 75 phylogenetically diverse large subunit ribosomal RNA sequences was created and searched for secondary and tertiary structure elements by computer. The search revealed four unknown secondary structural pairings, two internal loop closings, and five short-range tertiary interactions—two of which were pairings of an unusual type. One brings a loop together with two other loops previously known to be paired, and one involves a nucleotide within a presumed tetraloop. The latter interaction constrains the RNA structure near the ribosomal E-site, where two base pairs previously suggested to be in parallel orientation are now proven. No clear phylogenetic evidence for direct base pairing between the large and small subunit rRNA was found.**

Except for x-ray crystallography and NMR spectroscopy, comparative sequence analysis can resolve rRNA structure in greatest detail. Moreover, the structures identified by comparative analysis potentially include transient interactions that may occur at various stages in the ribosome cycle. The power of comparative analysis lies in the fact that only information from (properly) aligned sequences is used, rather than (imprecise) folding algorithms or experimental data. Its rationale is that, since specific structure–function constraints define evolutionarily permissible sequence variations, positions in a sequence alignment that clearly covary are likely to be in physical contact in the actual molecule; covariation of base pairs in known double helices is frequently seen. Covariation analysis has been used to demonstrate all secondary and many tertiary contacts in tRNA (1, 2). It also finds secondary structures in agreement with experimental data for 5S rRNA (e.g., ref. 3) and has been the primary instrument by which 16S rRNA (e.g., ref. 4) and 23S rRNA (e.g., refs. 5–8) secondary structure and tertiary contacts have been revealed.

In this study the significantly extended data base was searched for covariances by a dedicated computer approach to update the phylogenetically proven bacterial large subunit ribosomal RNA structure, to detect universal secondary and tertiary structure elements, and to gain a deeper sense of rRNA structures in general.

METHODS

Sequence Alignments. Alignments were created as described (8) using the sequence editor ALMA (9) on a DEC VAX/VMS computer.

General Guidelines. Since the criteria for inclusion of base pairs in any secondary structure model vary considerably, the basic rules and simple concepts used here will be stated. Positive evidence for a base–base interaction is given when the two positions in question covary in composition (from sequence to sequence) a significant number of phylogenetically independent times (within Watson–Crick pairing constraints for a base pair). Variation in composition of one position in the absence of variation in the other is taken as negative evidence, disfavoring or disproving the interaction. Those possible base pairs for which neither positive nor negative evidence exists are also excluded. A base pair is considered proven if there is at least twice as much positive evidence as negative. As a general rule, when the ratio is less, it is preferable not to include it. However, when a base pair is supported in one primary kingdom [now termed domain (10)] and disproven in others, it is considered specific for that one domain. If an interaction is supported by an inconclusive number of covariations, the following observations strengthen the case of its existence: Protection against chemical modification, occurrence of nearby crosslinks, structural proximity of the involved bases, covariations occurring between closely related species, and absence of an equally well supported alternative together with covariations occurring at neighboring positions. However, additional variations in additional sequences most often provide the needed positive or negative evidence.

Computer Implementation. The covariation search program CBCFOLD (unpublished) allows user-defined base pairs, helix pairing length, number of mismatches and covariations, rules for covariance, weighting covariations by an evolutionary distance matrix, etc. Assisted by a user-defined mask that marks unambiguous alignment, the program searches through all possible alignment alternatives. The output comprises alignment section files with the suggested pairings shown in reverse video when viewed at an ordinary text terminal. CBCFOLD is a Fortran program made for DEC VAX/VMS computers.

RESULTS

Sequence Alignments. An alignment of 23S-like sequences from 27 bacteria, 7 chloroplasts, 15 archaea, 16 eukaryotes, and 10 mitochondria was created. (It contained all cytoplasmic sequences known at the time of analysis and has been recently updated.) In addition, alignments were made of 49 16S-like and 42 5S sequences that correspond to these sequences.

Secondary Structure. Fig. 1 shows, for *Escherichia coli* 23S rRNA, the secondary and higher-order interactions that are now considered proven.

Secondary structural elements not present in the previous (most recent) model (7) are the pairings 19a, 43a, 49a, and 69a, evidence for which is presented in Table 1. Pairing 43a has also been described in another context (12). Covariance support for pairing 49a is marginal, but it was included given the structural proximity of the two bases (see Fig. 1) and their inaccessibility to chemical and enzymatic probes (6). Pairing 69a (involving G1935 and C1962) lies in the vicinity of two chemically crosslinkable oligonucleotides [CC at positions 1941–42 with GC at positions 1964–65 (11)]. However, N3 (Watson–Crick pairing position) of C1962 is reactive to

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

*Present address: Department of Microbiology, University of Illinois, 131 Burrill Hall, 407 South Goodwin Avenue, Urbana, IL 61801.

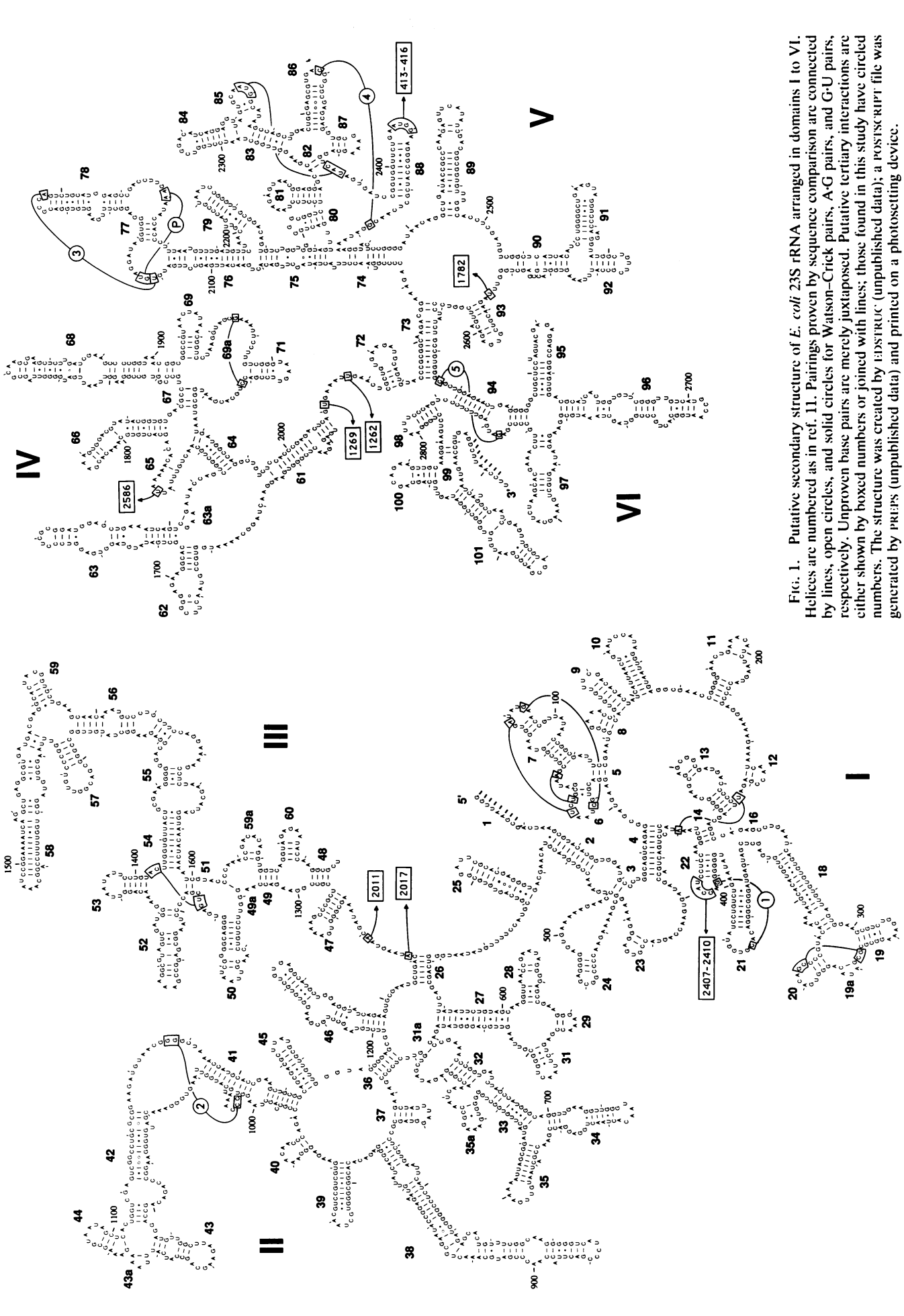

FIG. 1. Putative secondary structure of *E. coli* 23S rRNA arranged in domains I to VI. Helices are numbered as in ref. 11. Pairings proven by sequence comparison are connected by lines, open circles, and solid circles for Watson–Crick pairs, A·G pairs, and G·U pairs, respectively. Unproven base pairs are merely juxtaposed. Putative tertiary interactions are either shown by boxed numbers or joined with lines; those found in this study have circled numbers. The structure was created by EDSTRUC (unpublished data); a POSTSCRIPT file was generated by PREPS (unpublished data) and printed on a photosetting device.

Table 1. Listing of base pairs equivalent to the helical elements and tertiary interactions

		19a	43a	49a	69a	1	2	P	3	4	5
B	Thermotoga maritima	C G	U A	G C	G C	C G	CC GG	GU AA	U A	C G	C G
B	Herpetosiphon giganteus	C G	U A	G C	G C	C G	CC GG	GU AA	C G	C G	C G
*B	Chloroflexus aurantiacus		U A	G C	G C		CC GG	GU AA	C G	C G	C G
B	Deinonema *species*	C G	U A	G C	G C	C G	CC GG	GU AA	C G	C G	C G
B	Thermus thermophilus	C G	U A	G C	G C	C G	CC GG	GU AA	C G	C G	C G
B	Pirelula maritima	C G	U A	G -	G C	C G	CC GG	GU AA	U A	U A	C G
*B	Chlamydia psittaci	G C	U A	G C	G C	C G	NC GG	GU AA	U A	C G	C G
B	Taxeobacter ocellatus	C G	U A	G C	G C	A U	CC GG	GU AA	U A	C G	C G
B	Leptospira interrogans	C G	U A	G C	G C	C G	CC GG	GU AA	U A	C G	C G
*B	Chlorobium limicola	C G	U A	G C	G C	C A	CC GG	GU AA	U A	C G	C N
*B	Flexibacter flexilis	C G	U A	G C	G C	G C	CC GG	GU AA	U A	C G	C G
B	Flexithrix dorotheae	C G	U A	G C	G C	G C	CC GG	GU AA	U A	C G	C G
*B	Flexibacter sancti	C G		G C	G C	G C	CC	GU AA	U A	C G	C
*B	Flexibacter canadensis	C G		G C	G C	G C	CC	GU AA	U A	C G	C G
*B	Flexibacter aurantiacus	C G	U A	G C	G C	C G	CC GG	GU AA	U A	C G	
*B	Flavobacterium breve	C G		G C	G C	C G	CC GG	GU AA	U A	C G	
*B	Flavobacterium odoratum	C G	U A	G C	G C	C G	CC GG	GU AA	U A	C G	C N
*B	Campylobacter coli	C G	U A	G C	G C	C G	CC GG	GC AG	U A	C G	C G
*B	Flexibacter ruber	C G				G C					
*B	Flexibacter litoralis	C G				G C					
*B	Cytophaga diffluens	A U				G C					
*B	Flavobacterium meningosept.	C G				C G					
*B	Chlorobium vibrioforme	C G	U A	G C	G C		CC GG	GU AA	U A	C G	C G
B	Rhodobacter capsulatus	G -	U A	G C	G C	C G	CC GG	GC AG	U A	C G	U A
*B	Rhodobacter sphaeroides	C G	U A	G C	G C	C G	CC GG	GC AG	U A	C G	U A
*B	Agrobacterium tumefaciens	C G	U A	G C	G C	C G	CC GG	GU AA	U A	C G	U A
B	Pseudomonas cepacia	C G	U A	G C	G C	C G	CC GG	GU AA	U A	C G	C G
*B	Thiobacillus cuprinus	C G	U A	G C	G C	C G	CC GG	GU AA	U A	C G	C G
B	Pseudomonas aeroginosus	C G	U A	G C	G C	C G	CC GG	GU AA	U A	C G	C G
B	Ruminobacter amylophilus	G G	U A	G C	G C	C G	CC GG	GU AA	U A	C G	C G
B	Escherichia coli	G C	U A	G C	G C	C G	CC GG	GU AA	U A	C G	C G
*B	Heliobacterium chlorum	C G	U A	G C	G C	C G	CC GG	AC GG	U A	C G	C G
B	Micrococcus luteus	C G	U A	G C	G C	C G	CC GG	GU AA	U A	C G	C G
B	Streptomyces ambofaciens	C G	U A	G C	G C	C G	CC GG	GU AA	U A	C G	C G
B	Staphylococcus carnosus	C G	U A	G C	G C	C G	CC GG	AC GG	U A	C G	C G
B	Corynebacterium glutamicum	C G	U A	G C	G C	C G	CC GG	GU AA	U A	C G	C G
B	Mycoplasma pneumoniae	G A	U A	G C	G C	C G	CC GG	AG GA	U A	C G	C G
B	Mycoplasma hyopneumoniae	C G	U A	G C	G C	C G	CC GG	GU AA	U A	C G	C G
B	Lactobacillus lactis	C G	U A	G C	G C	C G	CC GG	AC GG	U A	C G	C G
B	Staphylococcus aureus	C G	U A	G C	G C	C G	CC GG	AC GG	U A	C G	C G
B	Streptococcus oralis	G	U A	G C	G C	C G	CC GG	AC GG	U A	C G	C G
B	Listeria monocytogenes	C G	U A	G C	G C	C G	CC GG	AC GG	U A	C G	C G
B	Bacillus subtilis	C G	U A	G C	G C	C G	CC GG	AC GG	U A	C G	C G
B	Bacillus globisporum	C G	U A	G C	G C	C G	CC GG	AC GG	U A	C G	C G
B	Bacillus stearothermophilus	C G	U A	G C	G C	C G	CC GG	AC GG	U A	C G	C G
B	Anacystis nidulans	C G	U A	G C	G C	C G	CC GG	GC AG	C G	C G	C G
C	Euglena gracilis	C G	U A	G C	G C	C G	CC GG	GC AG	C G	U A	C G
C	Zea mays	C G	U A	G C	G C	C G	CC GG	GC AG	C G	C G	C G
C	Oryza sativa	C G	U A	G C	G C	C G	CC GG	GC AG	C G	C G	C G
C	Nicotiana tabacum	C G	U A	G C	A C	C G	CC GG	GC AG	C G	C G	C G
C	Astasia longa	C G	U A	G C	G C	C G	CC GG	GC AG	C G	U A	C G
C	Marchantia polymorpha	C G	U A	G C	G C	C G	CC GG	GC AA	C G	C G	C G
C	Chlorella ellipsoidea	C G	U A	G C	G C	C G	CC GG	GC AG	C G	C G	C G
A	Pyrobaculum islandicum		C G	G C	C G	- -	NN GG	AG GA	C G	- -	C G
A	Thermoproteus tenax	C G	C G	G C	C G	- -	CC GG	AG GA	C G	- -	C G
A	Thermophilum pendens	C G	C G	G C	C G	- -	CC GG	GU AA	C G	- -	C G
A	Pyrodictium occultum			G C	C G	- -		GU AA	C G	- -	U A
A	Staphylothermus marinus			G C	C G	- -		GU AA	C G	- -	U A
A	Desulfurococcus mobilis	C G	U A	G C	C G	- -	CC GG	GU AA	C G	- -	U A
A	Sulfolobus solfataricus	C G	C G	G C	C G	- -	CC GG	GU AA	C G	- -	U A
A	Thermoplasma acidophilum	C G	U A	G C	G U	- -	CC GG	GU AA	C G	- -	U A
A	Thermococcus celer	C G	U A	G C	G C	- -	CC GG	GC AG	C G	- -	U A
A	Pyrococcus furiosus			G C	G C			GC AG	C G	- -	U A
*A	Archaeoglobus fulgidus	C G	U A	G C	G U	- -	CC GG	GC AG	C G	- -	U A
A	Halobacterium halobium	G C	C G	G C	G U	- -	CC GG	GC AA	U A	- -	U A
A	Halococcus morrhuae	G C	C G	G C	G U	- -	CC GG	GC AA	U A	- -	U A
A	Halobacterium marismortui	G C	C G	G C	G U	- -	CC GG	GC AA	U A	- -	U A
*A	Haloferax volcanii	G C	C G	G C	G U	- -	CC GG	GC AA	U A	- -	U A
*A	Microspirillum hungatii	C G	U A	G C	G U	- -	CC GG	GC AG	C G	- -	U A
A	Methanobacterium thermoaut.	C G	U A	G C	G C	- -	CC GG	GC AG	U A	- -	U A
A	Methanococcus vannielii	U G	U A	G C	G C	- -	CC GG	AC GG	U A	- -	U A
E	Physarum polycephalum	U G	C G	G C	C G	- -	CU AG	GU AA	U U	A A	- -
E	Tetrahymena thermophila	C G	C G	G C	C G	- -	GC GC	AU UA	U A	A U	- -
E	Prorocentrum micans	C G	C G	G C	C G	- -	GC GC	GU AA	U G	A U	- -
E	Mucor racemosus	C G	C G	A U	C G	- -	GC GC	GU AA	U G	G C	- -
E	Dictyostelium discoideum	C G	C G	G C	C G	- -	GU AC	GU AA	U A	A U	- -
E	Saccharomyces cerevisiae	C G	C G	A U	C G	- -	GC GC	GU AA	U G	G C	- -
E	Citrus limon	C G	C G	A U	C G	- -	GC GC	GU AA	U A	A U	- -
E	Oryza sativa	C G	C G	A U	C G	- -	GC GC	GU AA	U G	A U	- -
E	Lycopersicon esculentum	C G	C G	A U	C G	- -	GC GC	GU AA	U A	A U	- -
E	Caenorhabditis elegans	C G	C G	G C	C G	- -	GC GC	GU AA	U G	A U	- -
E	Xenopus laevis	C G	C G	G C	C G	- -	GC GC	GU AA	U G	A U	- -
E	Drosophila melanogaster	G A	C G	G C	C G	- -	CC GC	GU AA	U G	A U	- -
E	Crithidia fasciculata	C G	C G	G U	C G	- -	GU AC	GU AA	U A	A U	- -
E	Rattus norvegicus	C G	C G	G C	C G	- -	GC GC	GU AA	U -	A U	- -
E	Homo sapiens	C G	C G	G C	C G	- -	GC GC	GU AA	U -	A U	- -
E	Mus musculus	C G	C G	G C	C G	- -	GC GC	GU AA	U -	A U	- -
M	Paramecium tetraurelia	U G	U A	G -	A C	- -	GC GC	AA AA	U -	U G	- -
M	Oenothera berteriana	C G	U A	G C	G C	- -	CC GG	GU AA	U -	U G	- -
M	Zea mays	C G	U A	G C	G C	- -	CC GG	GU AA	U -	U G	- -
M	Saccharomyces cerevisiae	- -	U U	G U	G C	- -	UC AA	AU AG	U -	G G	- -
M	Anacystis nidulans	- -	U A	G C	G C	- -	UC CA	-- --	- -	A A	- -
m	Homo sapiens	- -	A A	- -	A C	- -	-- --	-- --	- -	- -	- -
m	Bos taurus	- -	A A	- -	A U	- -	-- --	-- --	- -	- -	- -
m	Mus musculus	- -	A A	- -	A U	- -	-- --	-- --	- -	- -	- -
m	Xenopus laevis	- -	G A	- -	G U	- -	-- --	-- --	- -	- -	- -
m	Aedes albopictus	- -	- -	- -	A U	- -	-- --	-- --	- -	- -	- -

Pairings are numbered as in Fig. 1. Dashes indicate sequence not present or not unambiguously aligned; blanks represent incomplete sequence. Organisms are grouped as bacteria (B), chloroplasts (C), archaea (A), eukaryotes (E), and

Table 2. Accessibilities to chemical and enzymatic probes

	Pairings																					
	19a		43a		49a		69a		1		2			P			3		4		5	
Isolated RNA																						
DMS	–	–	–	–	–	–	–	–	–	–	–	–	–	–	W	W	–	–	–	–	–	–
DEP	–	–	–	–	–	–	–	–	–	–	–	–	W	–	–	–	–	–	–	–	–	–
Kethoxal	–	–	–	–	–	–	W	–	–	W	–	–	W	W	S	–	–	–	–	–	–	S	–	S
CMCT	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
RNase T1	–	–	–	–	–	–	W	–	–	–	–	–	–	–	–	–	–	–	–	–	–	W	–	S
RNase T2	–	–	–	–	–	–	–	–	–	–	S	–	–	–	–	–	–	–	–	–	W	–	–	–
RNase CV	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
rRNA–protein complex																						
DMS	–	–	–	–	–	–	–	S	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
DEP	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Kethoxal	–	–	–	–	–	–	–	–	–	–	–	–	–	–	S	–	–	–	–	–	–	–	–	–
CMCT	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
RNase T1	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
RNase T2	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
RNase CV	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Cross-link	–		–		–		+		–		–			–			+		–		–	

Pairings are numbered as in Table 1. Data for isolated RNA and the RNA–protein complex are shown (5). Dashes indicate protection; S and W denote weak and strong reactivity, respectively; a plus sign indicates a known crosslink at, or close to, the involved nucleotides (10). DMS, dimethyl sulfate; DEP, diethyl pyrocarbonate; CMCT, carbodiimide metho-*p*-toluene sulfonate.

dimethyl sulfate in intact 50S subunits, suggesting that the pairing may not occur under all conditions.

Additional pairings within internal loops now considered proven are C1800·G1817 in helix 66, U1856·U1886 and A1858·G1884 in helix 68, C2521·G2544 in helix 91, and UGG (positions 2684–86)·CUA (positions 2725–23) within helix 96. Although covariations exist for the pairing C2789·G2892, negative evidence also exists.

The following are minor differences between this model and its immediate predecessor (7). The highly conserved helix 70, which was unproven but previously suggested because it had base pairing potential, is now disproven and excluded. Also disproven are the internal loop closings in helices 7, 11, 62, and 91 and two base pairs at the junction of helices 75, 76, and 79. In addition, the present diagram eliminates a number of unproven base pairs at the ends of helices. Additional sequences and use of common rules for evaluation could eliminate these detailed discrepancies in the future.

Tertiary Interactions. An exhaustive CBCFOLD search was made, in which all unpaired regions of the nonmitochondrial sequences were intercompared to locate additional base–base interactions. All published tertiary interactions (7, 8) were detected and some were confirmed by covariations in sequences not previously available. The one exception, excluded from the present model, is the U2016·A2058 interaction [suggested earlier (8) but for which evidence is now conflicting].

Five tertiary interactions for which good evidence exists were also found. They are connected by lines with encircled numbers 1–5 in Fig. 1, and their supporting evidence is listed in Table 1. Pairings 1, 3, and 5 derive further support from unpublished or recently sequenced RNAs (kindly provided by C. R. Woese and W. Ludwig, personal communication). The accessibilities of the involved nucleotides to chemical and enzymatic probes are tabulated in Table 2.

In addition to the interactions listed in Table 1, others were found for which the evidence was both positive and negative. Also, a number of pairings were found for which the evidence was exclusively positive but weak; an example is U2739·A2749, whose composition is constant in archaea, eukaryotes, and bacteria, except for *Herpetosiphon giganteus*, which exhibits a G·C pair.

Contacts with Other rRNAs. To check whether base pairing occurs between the two ribosomal subunits, alignments of corresponding large and small subunit rRNA were searched. No proven pairings were found. Therefore, either there are none or they occur in highly conserved regions. As judged from image averaging contours of the ribosomal subunits (13), the contacting surfaces are minimal.

The aligned large subunit ribosomal RNAs were similarly searched against the corresponding 5S rRNAs. No interactions considered proven emerged. Although covariations among bacteria suggest a pairing between U2477 and 5S A109, additional supporting evidence must be found before this putative interaction is accepted.

DISCUSSION

The described tertiary interactions constrain local RNA folding in interesting ways. Pairings 2 and 5 (Fig. 1) are typical "pseudoknots" (14, 15), as are a number of the other known tertiary interactions. It is becoming clear that such local back-foldings occur frequently in rRNA (additional possible short-range pseudoknots are also seen among the insufficiently proven pairings). However, pairings 1 (C385·G411), P (G2112·A2169 and U2113·A2170), and 3 (U2111·A2147) are unlike any previously known pairings, to our knowledge, and are discussed separately below.

The C385·G411 pairing brings the terminal loop of helix 21 together with the loops of helices 22 and 88, which were known to be paired (8). This leaves only A412, within the loop of helix 22, unpaired. It is stereochemically feasible to extend helix 22 coaxially by the long-range interaction between its terminal loop and that of helix 88. If so, a sharp turn must then occur at A412, exposing this nucleotide and rendering it reactive to chemical modification. Indeed, chemical modification experiments are in agreement with this prediction: In

mitochondria from lower (M) and higher (m) eukaryotes, and within these groups organisms are in approximate phylogenetic arrangement (i.e., related species are adjacent).

*Unpublished or recently sequenced RNAs that were not used in the analysis.

intact 50S subunits, all nucleotides within the loop of helix 22 are protected except for A412, the dimethyl sulfate reactivity of which is enhanced relative to isolated RNA (6).

The pairings G2112·A2169 and U2113·A2170, which have been described as parallel base pairs (7), are here confirmed by additional evidence (see columns P in Table 1). The G2112·A2169 pair may not involve N1 and N2 of G2112, since these positions are strongly reactive with kethoxal in isolated RNA and 50S subunits (6). These pairings [which occur near the ribosomal exit site (16)] are the only known parallel base pairs in rRNA.

The pairing U2111·A2147 positions the terminal loop of helix 78 [which binds ribosomal protein L1 (17)] adjacent to the above mentioned parallel interaction [both U2111 and A2147 are protected against chemical modification under all conditions tested (6)]. The terminal loop can be crosslinked to the region from position 2201 to position 2204 (11), indicating that this region could also be close to the parallel pairings G2112·A2169 and U2113·A2170 [which lie near the ribosomal E-site (16)]. In addition to position U2111, position A2147 appears to covary with position G2144 (data not shown) in a manner typical for "tetraloops" (18), in which the two terminal bases in the loop (in this case positions G2144 and G2147) interact with one another (19). There are two alternative explanations for the observed covariations between position A2147 and two other positions: N3 and N2 in the ring of A2147 could be involved in a triple base pair that includes U2111·A2147 (Watson–Crick pairing) and G2144. Alternatively, the two implied pairings of A2147 need not occur simultaneously. If the latter is correct, then this would be an example of transient interactions that form only at particular stages in the translational cycle.

The structure of a number of elements in the large subunit ribosomal RNA has now been highly constrained by covariance analysis, making them excellent candidates for more extensive analysis by physical-chemical techniques. Several such regions that might be crystallized (in isolation) are G54 → C116, U290 → G350, G1051 → U1108, and U2105 → A2184 (see Fig. 1).

Eventually, results obtained by sequence comparison will help to trace the RNA chains within x-ray electron density maps of intact 50S ribosomal subunits (13). Even more importantly, refined comparative sequence analyses could in principle be used to detect functional details and alternative functional conformations that are otherwise difficult or impossible to obtain, even by x-ray crystallography.

I am grateful to Jens Nyborg for support and to Carl Woese and Wolfgang Ludwig, who allowed me to use their sequences prior to publication. I also thank Erko Stackebrandt for sending small subunit rRNA sequences. This work was supported by the Bioregulation Research Center, Århus University, under the Danish Biotechnology Program.

1. Levitt, M. (1969) *Nature (London)* **224,** 759–763.
2. Olsen, G. (1983) Ph.D. dissertation (Univ. of Colorado, Denver).
3. Fox, G. E. & Woese, C. R. (1975) *Nature (London)* **256,** 505–507.
4. Woese, C. R., Gutell, R. R., Gupta, R. & Noller, H. F. (1987) *Microbiol. Rev.* **47,** 621–669.
5. Gutell, R. R. & Fox, G. E. (1988) *Nucleic Acids Res.* **16,** Suppl., r175–r269.
6. Egebjerg, J., Larsen, N. & Garrett, R. A. (1990) in *The Ribosome: Structure, Function and Evolution*, eds. Hill, W., Dahlberg, A., Garrett, R. A., Moore, P. B., Schlessinger, J. & Warner, J. R. (Am. Soc. Microbiol., Washington), pp. 168–179.
7. Gutell, R. R. & Woese, C. R. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 663–667.
8. Leffers, H., Kjems, J., Østergaard, L., Larsen, N. & Garrett, R. A. (1987) *J. Mol. Biol.* **195,** 43–61.
9. Thirup, S. & Larsen, N. (1990) *Proteins Struc. Func. Genet.* **7,** 291–295.
10. Woese, C. R., Kandler, O. & Wheelis, M. L. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 4576–4579.
11. Mitchell, P., Osswald, M., Schueler, D. & Brimacombe, R. (1990) *Nucleic Acids Res.* **18,** 4325–4333.
12. Egebjerg, J., Douthwaite, S. R., Liljas, A. & Garrett, R. A. (1990) *J. Mol. Biol.* **213,** 275–288.
13. Yonath, A., Bennett, W., Weinstein, S. & Wittmann, H. G. (1990) in *The Ribosome: Structure, Function and Evolution*, eds. Hill, W., Dahlberg, A., Garrett, R. A., Moore, P. B., Schlessinger, J. & Warner, J. R. (Am. Soc. Microbiol., Washington), pp. 134–147.
14. Pleij, C. W. A., Rietveld, K. & Bosch, L. (1985) *Nucleic Acids Res.* **13,** 1718–1731.
15. Puglisi, J. D., Wyatt, J. R. & Tinoco, I., Jr. (1990) *J. Mol. Biol.* **214,** 437–453.
16. Moazed, D. & Noller, H. F. (1989) *Cell* **57,** 585–591.
17. Branlant, C., Krol, A., Sriwidada, J., Ebel, J. P., Sloof, P. & Garrett, R. A. (1976) *Eur. J. Biochem.* **70,** 457–469.
18. Woese, C. R., Winker, S. & Gutell, R. R. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 8467–8471.
19. Cheong, C., Varani, G. & Tinoco, I., Jr. (1990) *Nature (London)* **346,** 680–682.