

Collection of small subunit (16S- and 16S-like) ribosomal RNA structures: 1994

Robin R. Gutell

MCB Biology, Campus Box 347, University of Colorado, Boulder, CO 80309-0347, USA

ABSTRACT

A collection of diverse 16S and 16S-like rRNA secondary structure diagrams are available. This set of rRNAs contains representative structures from all of the major phylogenetic groupings — *Archaea*, *(eu)Bacteria*, and the nucleus, mitochondrion, and chloroplast of *Eucarya*. Within this broad phylogenetic sampling are examples of the major forms of structural diversity currently known for this class of rRNAs. These structure diagrams are available online through our computer-network WWW server and anonymous ftp, as well as from the author in hardcopy format.

INTRODUCTION

Comparative structure analysis of 16S rRNA has progressed for the past 15 years. Initially with a minimal number of sequences and a simpler perception of comparative analysis, a secondary structure common to the few 16S and 16S-like rRNA sequences was proposed (reviewed in 1,2.). Since then, the number of publicly available and complete (or nearly complete) 16S and 16S-like rRNA sequences has increased most dramatically. As of June 1994, this number is approximately 3,100 (see Table 1; Ribosomal RNA Database Project — see article in this issue; RRG — unpublished collection). In parallel with this pronounced increase in 16S rRNA sequences, our comparative methods and interpretation have evolved, resulting in a more detailed 16S rRNA secondary and tertiary structure (reviewed in 1).

The details of the current secondary and tertiary structure for 16S rRNA should not be considered final. While the vast majority of the proposed secondary and tertiary structure interactions are considered correct given their high degree of comparative evidence, alterations in these structural elements are not anticipated. However, a small number of proposed interactions are considered tentative, for these contain a minimal amount of comparative evidence and are thus candidates for secondary structure adjustments in the future. Further refinement in 16S rRNA higher-order structure will also result from the analysis of a larger number of aligned sequences with improved correlation analysis methods (3; Gautheret, Damberger, and Gutell, mss. in preparation). These latter structural constraints are likely to involve base triples, non-canonical pairings, and other structural elements more complex than *simple* secondary structure base pairings.

OBJECTIVES

The primary goals for this 16S rRNA database are to: (1) Present the most recent comparatively inferred secondary and tertiary structure for 16S rRNA. Minor refinements in these structures are expected, as we identify new (or alternative) secondary and tertiary interactions. (2) Offer a sampling of representative 16S and 16S-like secondary structure diagrams. The initial selection of structures are phylogenetically as well as structurally distinct from one another (see Table 2). Over time additional examples of diverse structures will be generated, as well as structures that are closely related to members of this database. (3) Three divergent 16S-like rRNA diagrams are presented here. They are: *Haloferax volcanii*, a representative *Archaea* (5, Figure 1); *Phreatamoeba balamuthi*, a lower *Eucarya* that contains several large insertions (6, Figure 2); and the *Chlamydomonas reinhardtii* mitochondrion; a most bizarre rRNA that is coded for by a fragmented rRNA gene (7, Figure 3).

DATABASE

The higher-order structures described herein will be available on-line or in hardcopy format from the author. On-line access is by anonymous ftp or through the WWW (World Wide Web). These computer files will be distributed in PostScript format only. Our ftp address and directory are:

pundit.colorado.edu (128.138.212.53)
/pub/RNA/16S

Our WWW server can be accessed by programs like *XMOSAIC* (see Damberger and Gutell, this issue), that navigate the WWW. With this system the secondary structures can be viewed and

Table 1. Phylogenetic distribution of publicly available 16S and 16S-like rRNA sequences (as of June 1994)

Organism type	# 16S rRNA
Archaea	120
(Eu)Bacteria	2130
Eucarya (nuclear)	700
Organelles	
Mitochondrion	110
Plastid	40
TOTAL	3100

Table 2. List of currently available 16S and 16S-like rRNA secondary structure diagrams. Phylogenetic classification is based on Woese *et al.* (4)

Archaea	
Euryarchaeota	<i>Archaeoglobus fulgidus</i> , <i>Halobacterium marismortui</i> (HC8), <i>Halobacterium marismortui</i> (HC10), <i>Haloferax volcanii</i> , <i>Methanobacterium formicicum</i> , <i>Methanospirillum hungatei</i> , <i>Methanococcus vannielii</i> , <i>Thermococcus celer</i> , <i>Thermoplasma acidophilum</i>
Crenarchaeota	<i>Pyrodicticum occultum</i> , <i>Sulfolobus solfataricus</i> , <i>Thermoproteus tenax</i>
(eu)Bacteria	
Aquificales	<i>Aquifex pyrophilus</i>
Thermotoga	<i>Thermotoga maritima</i> , <i>Geotoga subterranea</i> , <i>Petrotoga miotherma</i>
Deinococcus and relatives	<i>Thermus thermophilus</i>
Spirochetes and relatives	<i>Borrelia burgdorferi</i> , <i>Leptonema illini</i> , <i>Spirochaeta aurantia</i>
Planctomyces and relatives	<i>Planctomyces staleyi</i> , <i>Gemmata obscuriglobus-1</i> , <i>Gemmata obscuriglobus-2</i>
Green sulfur bacteria	<i>Chlorobium vibrioforme</i>
Green non-sulfur bacter.	<i>Thermomicrobium roseum</i>
Cytophaga-Flexi.-Bacter.	<i>Bacteroides fragilis</i>
Purple bacteria	<i>Agrobacterium tumefaciens</i> , <i>Pseudomonas testosteroni</i> , <i>Escherichia coli</i> , <i>Dichelobacter nodosus</i> , <i>Myxococcus xanthus</i> , <i>Desulfovibrio desulfuricans</i> , <i>Campylobacter sputorum</i> subsp. <i>sputorum</i>
Cyanobacteria	<i>Synechococcus</i> sp. 6301
Gram-positive	<i>Helibacterium chlorum</i> , <i>Arthrobacter globiformis</i> , <i>Frankia</i> sp., <i>Streptomyces coelicolor</i> , <i>Bacillus subtilis</i> , <i>Epulopiscium</i> sp. (strain A1), <i>Clostridium innocuum</i> , <i>Mycoplasma hyopneumoniae</i> , <i>Mycoplasma gallisepticum</i> , <i>Mycoplasma capricolum</i> ,
Eucarya (nuclear)	
Archezoa	<i>Hexamita</i> , <i>Giardia muris</i> , <i>Giardia ardeae</i> , <i>Giardia intestinalis</i> , <i>Encephalitozoon cuniculi</i> , <i>Vairimorpha necatrix</i>
Protoctista	<i>Phreatamoeba balamuthi</i> , <i>Babesia bigemina</i> , <i>Chlorella luteoviridis</i> , <i>Staurastrum</i> - M752, <i>Genicularia spirotaenia</i> , <i>Gracilariaopsis</i> sp., <i>Tritrichomonas foetus</i>
Fungi	<i>Saccharomyces cerevisiae</i> , <i>Cryptococcus neoformans</i>
Plantae	<i>Fragaria ananassa</i> (strawberry)
Animalia	<i>Drosophila melanogaster</i> , <i>Placopecten magellanicus</i> , <i>Xenopus laevis</i> , <i>Mus musculus</i> , <i>Homo sapiens</i>
Mitochondrion	
Protoctista	<i>Chlamydomonas reinhardtii</i> , <i>Paramecium tetraulia</i> , <i>Tetrahymena pyriformis</i> , <i>Physarum polycephalum</i>
Fungi	<i>Saccharomyces cerevisiae</i> , <i>Aspergillus nidulans</i>
Plantae	<i>Zea mays</i> (maize)
Animalia	<i>Bos taurus</i> (bovine), <i>Ascaris suum</i> , <i>Caenorhabditis elegans</i>
Chloroplast	
Protoctista	<i>Chlamydomonas reinhardtii</i> , <i>Chlorella vulgaris</i> , <i>Cryptomonas</i> sp., <i>Olisthodiscus luteus</i> , <i>Astasia longa</i> , <i>Euglena gracilis</i> , <i>Pylaiella littoralis</i> , <i>Cyanidium caldarium</i> , <i>Palmaria palmata</i>
Plantae	<i>Nicotiana tabacum</i> (tobacco), <i>Zea mays</i> (maize), <i>Marchantia polymorpha</i> (liverwort)

downloaded, while ftp access only supports downloading of the PostScript files. Our WWW URL address is:

URL:<http://pundit.colorado.edu:8080/RNA/16S/16s.html>

The work presented here complements the effort of the RDP (Ribosomal Database Project, see this issue); Their objectives are to generate and present various types of ribosomal RNA data and interpretation. On-line access to the 16S rRNA secondary structure PostScript files is also available from the RDP

anonymous ftp site (rdp.life.uiuc.edu). Additional information about this server and access to the Ribosomal Database Project is described in this special Database issue.

The interactive graphics program XRNA (developed by Bryn Weiser, UC Santa Cruz) facilitated the generation of the higher-order structure diagrams. Manual alignment of these 16S and 16S-like rRNA sequences was facilitated with the interactive sequence alignment editor AE2 (developed by Tom Macke, Scripps Clinic).

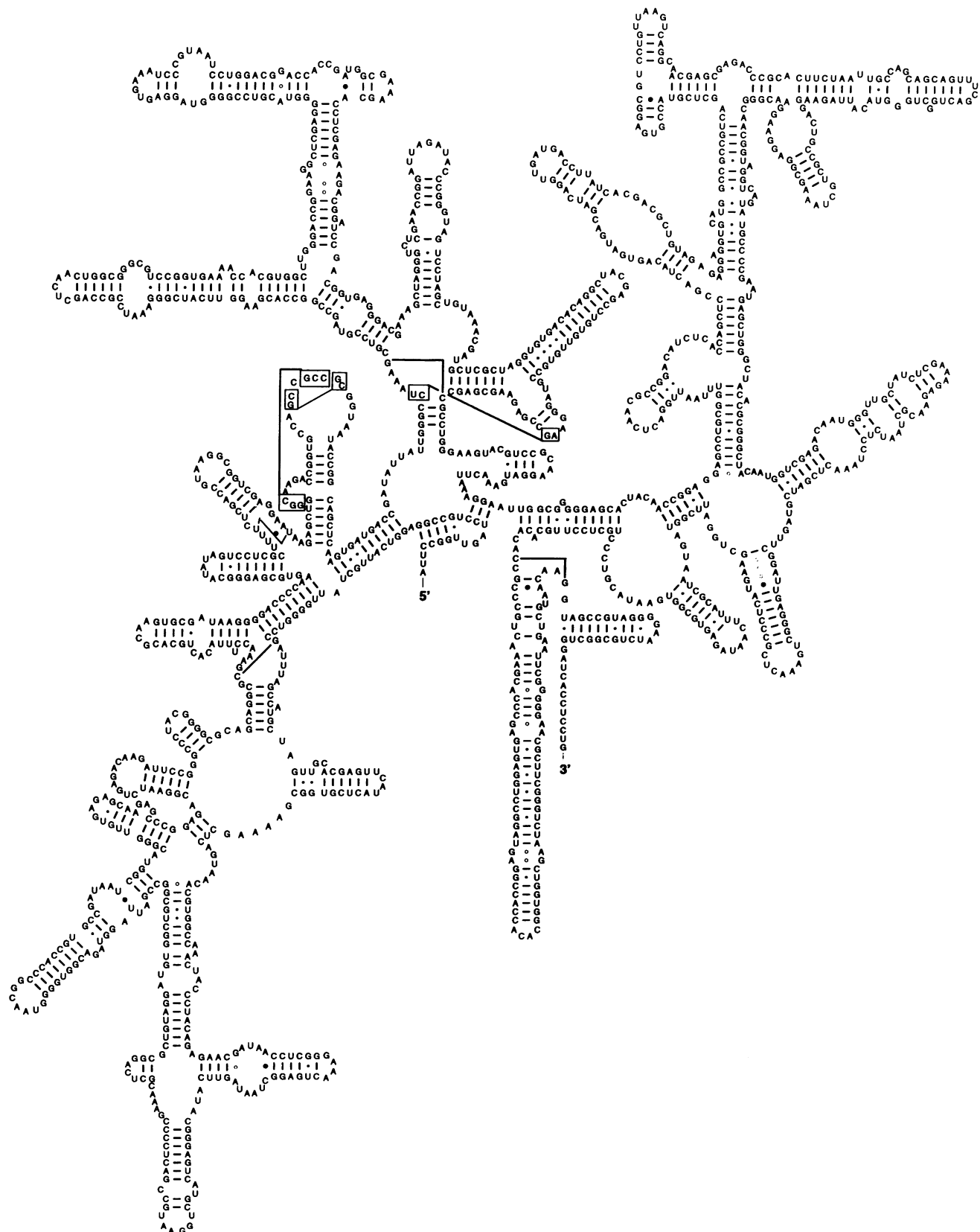


Figure 1. Higher-order structure diagram for *Haloferax volcanii* 16S rRNA (5). Different types of secondary structure basepairings are distinguished as follows: short lines connect canonical pairs (C:G, U:A), G:U pairs are denoted with larger open circles, and other non-canonical pairings with closed circles. Tertiary interactions are connected with thicker and longer solid lines. [Sequence Accession number is K00421].

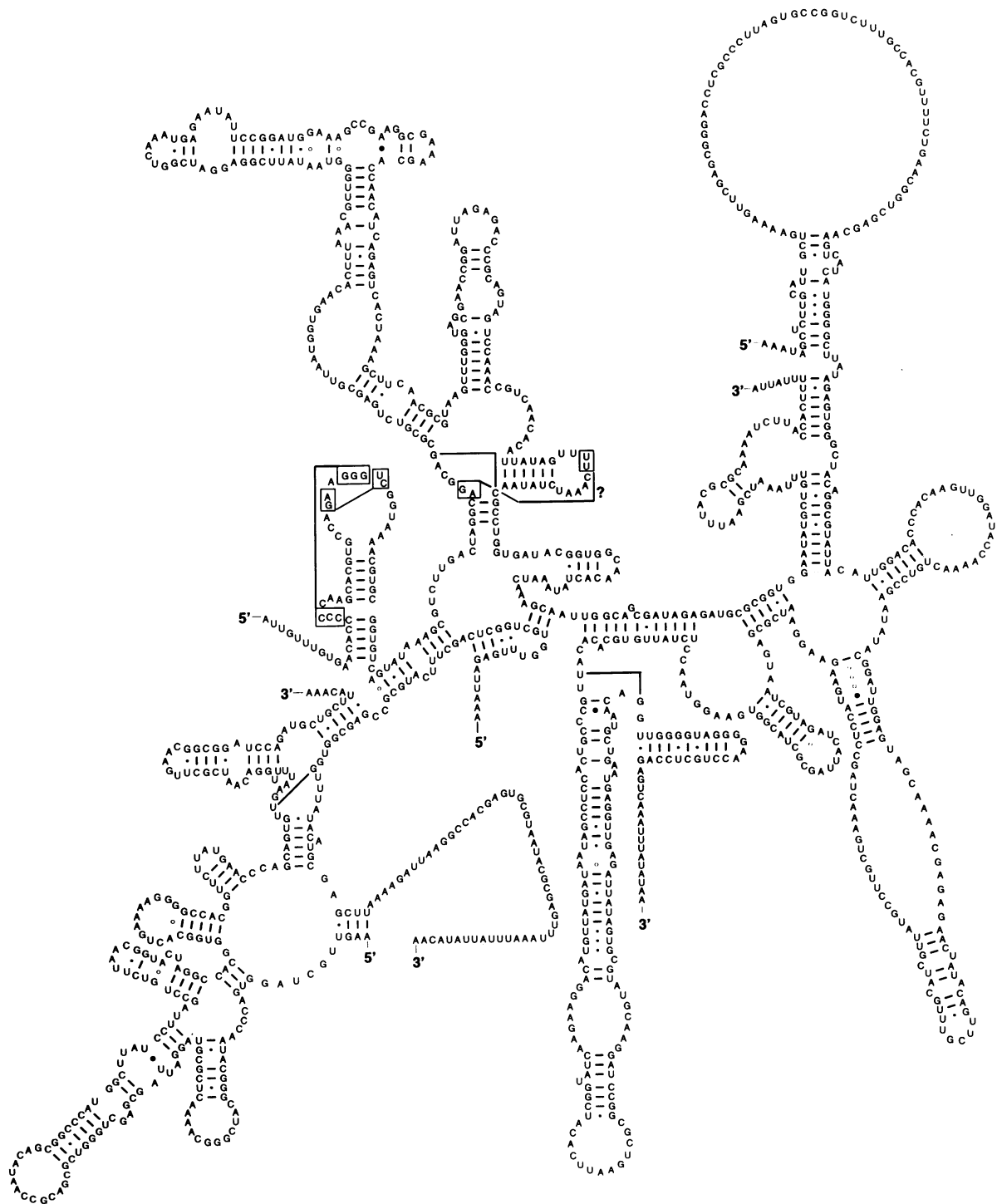


Figure 3. Higher-order structure diagram for *Chlamydomonas reinhardtii* mitochondrial 16S-like rRNA (7). Higher-order structure interactions illustrated and noted as in Fig. 1. [Sequence Accession number is X54860].

Figure 2. Higher-order structure diagram for the *Eucarya*, *Phreatamoeba balamuthi* nucleocytoplasmic 16S-like rRNA (6). Higher-order structure interactions illustrated and noted as in Fig. 1. [Sequence accession number is L23799].

Research benefiting from the secondary structure diagrams in hardcopy format or from the online compendium should cite this article.

ACKNOWLEDGEMENTS

I wish to thank Mr Simon Damberger for helping establish our WWW server, the W.M.Keck Foundation for their generous support of RNA science on the Boulder campus, and *SUN Microsystems* for their timely donation of computer equipment. The author is an Associate in the Program in Evolutionary Biology of the *Canadian Institute of Advanced Research*. This work was supported by the NIH (GM 48207).

REFERENCES

1. Gutell R.R., Larsen N., Woese C.R. (1994). Lessons from an Evolving Ribosomal RNA: 16S and 23S rRNA Structure from a Comparative Perspective. *Microbiological Reviews*. 58,10–26.
2. Gutell R.R. (1993). Comparative Studies of RNA: Inferring Higher-Order Structure from Patterns of Sequence Variation. *Current Opinion in Structural Biology*. Section: Nucleic Acids. 3,313–322.
3. Gutell R.R., Power A., Hertz G.J., Putz E.J., and Stormo G.D. (1992). *Nucleic Acids Res.* 20,5785–5795.
4. Woese C.R., Kandler O., and Wheelis M.L. (1990). Towards a natural system of organisms: Proposal for the domains *Archaea*, *Bacteria*, and *Eucarya*. *Proc. Natl. Acad. Sci. USA* 87,4576–4579.
5. Gupta, R., Lanter, J.M. and Woese, C.R. (1983). Sequence of the 16S ribosomal RNA from *Halobacterium volcanii*, an archaebacterium. *Science* 221, 656–659.
6. Hinkle, G., Leipe, D.D., Nerad, T.A. and Sogin, M.L. (1994). The unusually long small subunit ribosomal RNA of *Phreatamoeba balamuthi*. *Nucleic Acids Res.* 22, 465–469.
7. Boer, P.H. and Gray, M.W. (1988). Scrambled ribosomal RNA gene pieces in *Chlamydomonas reinhardtii* mitochondrial DNA. *Cell* 55, 399–411.