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

Robin Ray Gutell

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

INTRODUCTION

Inferring higher-order structure for complex RNA molecules, such as the ribosomal RNAs has relied primarily on comparative methods. Underlying these methods is the premise that molecules with different primary structure and similar functional characteristics have similar secondary and tertiary structure [Reviewed in: 1]. For these methods to be effective, the RNA molecules under study need to be sufficiently similar at the primary structure level to obtain good sequence alignments, however these same sequences also need to be proportionately different for positional covariance to occur, the indicator suggesting the existence of a basepair.

The higher-order structure models for 16S rRNA have evolved in stages. Initially, with a small number of 16S rRNA sequences in hand, a minimal secondary structure was proposed. Increases in the number and diversity of available 16S rRNA sequences and paralleled with improvements in correlation analysis algorithms has lead to the continual refinement of this structure model. In the earlier stages only secondary structure pairings were identified. In contrast during the latter stages only minor refinements in these pairings occurred while several novel tertiary and non-canonical pairing constraints were proposed [reviewed in: 2–3]. And now with over 2,200 16S and 16S-like rRNA available sequences spanning the three phylogenetic domains and the two organelles (Mitochondria and Chloroplast), detailed phylogenetic, structural, and structural evolution information is now being deciphered in great detail, although the resulting analysis from different groups is not always congruent.

Over the years several groups have developed 16S rRNA secondary structure models. The current versions for each are fairly analogous with one another [2, 4–7], although this has not always been the case. And while the current differences are small, some are significant for the *Escherichia coli* and other *Bacterial*, *Archaea*, *Eucarya*, and mitochondrial structure models. These versions should also not be considered final as these models are expected to undergo minor revisions as the number and diversity of 16S and 16S-like sequences increases and is paralleled with continued improvements and alternative correlation analysis interpretations [8, unpublished work].

Over the next few years this collection, and the accompanying 23S rRNA (see Gutell, Gray, Schnare, this issue) will grow in size and detail. The complexity of structure and the evolutionary dimension of these structures presents us with a wonderful opportunity to investigate RNA structural motifs and map with some precision, the evolution of these RNAs and underlying RNA structural characteristics associated with different phylogenetic assemblages. This collection of structures should also be of value

to the experimentalist studying rRNA structure and function. And equally valuable to those studying rRNA based phylogeny.

OBJECTIVES

The objectives for this annual collection are:

1. Present our most current comparative interpretation of the *Escherichia coli* 16S rRNA secondary structure [with general agreement between C.R. Woese, H.F. Noller, and myself]. While this secondary structure model is stable, minor adjustments are expected in some of the helical pairings.
2. Present other significant correlations that suggest tertiary and other complex structure in the *Escherichia coli* model. The numbers for such structural interactions should increase over the next few years.
3. Present a sampling of different 16S and 16S-like rRNA higher-order structure models. Starting with a broad sampling of phylogenetically and structurally distinct models, additional examples will be developed to fill in this broad phylogenetic and structure matrix. A parallel effort will refine all previously proposed models as new comparative structure information is obtained. To be discussed elsewhere and in more detail, this collection of structures will serve as the database for detailed analysis of RNA evolution and RNA structural motifs.
The first 16S and 16S-like structure release is set for summer/fall—1993, and will include major representatives from the three phylogenetic domains and organelles. The approximate numbers will be: 25 (*eu*)*bacteria*, 10 *archaea*, 15 *eucarya*, 2 chloroplast, and 10 mitochondria. Readers are encouraged to contact the author with suggestions for additional higherorder structure models not currently available.
4. Illustrated in this article are three divergent 16S and 16S-like rRNAs structure models for: *Escherichia coli*, a member of the (*eu*)*bacteria* phylogenetic domain; nuclear encoded *Saccharomyces cerevisiae*, a member of the *eucarya* domain; and *Caenorhabditis elegans* mitochondria, one of the smallest known 16S-like rRNAs.

This work is an adjunct to the monumental effort of the RDP (Ribosomal Database Project, see this issue); Their mandate being to generate and make various types of ribosomal RNA data and interpretation generally available. On-line access to these 16S rRNA secondary structure files is available from the RDP

Figure 1. Higher-order structure diagram for *Escherichia coli* 16S rRNA [2-3]. For secondary structure basepairings, short lines connect canonical pairs (C:G, U:A), G:U pairs are denoted with dots, A:G pairings with larger open circles, and other non-canonical pairings are connected with thicker and longer solid lines. Tertiary interactions are connected with thicker and longer solid lines. Every 10th nucleotide position is marked with a tic mark, while every 50th position is numbered [Sequence Accession number is J01695].

Figure 2. Higher-order structure diagram for *Saccharomyces cerevisiae* nucleocytoplasmic 16S-like rRNA. Higher-order structure interactions illustrated and noted as in Fig. 1. [Sequence is slightly modified from accession number J01353].

Figure 3. Higher-order structure diagram for *Caenorhabditis elegans* mitochondrial 16S-like rRNA. Higher-order structure interactions illustrated and noted as in Fig. 1. [Sequence Accession number is X54252].

anonymous ftp site (info.mcs.anl.gov). Currently, these PostScript files are maintained in the '/pub/RDP/SSU_rRNA/sec_struct' directory. Additional information about this server and access to the Ribosomal Database Project is described in their article in this issue. Readers unable to access this on-line information, can obtain a set of 16S rRNA hardcopy diagrams from the author.

The interactive graphics program *XRNA* (developed by Bryn Weiser, UC Santa Cruz) facilitated the generation of the higher-order structure models displayed here and the PostScript files available at the RDP anonymous ftp site. 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).

ACKNOWLEDGMENTS

Refining and interpreting these rRNA structure models is an ongoing and long term collaboration with Drs Carl Woese and Harry Noller. Bryn Weiser and Tom Macke are greatly acknowledged for developing the wonderful programs that make much of this analysis and presentation possible. I also wish to thank 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

- Gutell R.R. (1993a) *Curr. Opin. Struct. Biol.* 3, *in press*.
- Gutell R.R., Larsen N., and Woese C.R. (1993). In Zimmermann R.A., and Dahlberg A.E.(eds). *Ribosomal RNA: Structure, Evolution, Gene Expression and Function in Protein Synthesis*. CRC Press, Boca Raton, FL. *in press*.
- Gutell R.R. (1993b). In Nierhaus K.H., Subramanian A.R., Erdmann V.A., Franceschi F., and Wittman-Liebold B. (eds). *The Translational Apparatus*. Plenum Publishing Corporation, NY, NY *in press*.
- Brimacombe R., Greuer B., Mitchell P., Osswald M., Rinke-Appel J., Schuler D., and Stade K. (1990). In Hill W.E., Dahlberg A., Garrett R.A., Moore P.B., Schlessinger D., and Warner J.R. (eds). *The Ribosome: Structure, Function, and Evolution*. American Society for Microbiology, Washington D.C. pp93-106.
- Ehresmann B., Ehresmann C., Romby P., Mougél M., Baudin F., Westhof E., and Ebel J-P. (1990). In Hill W.E., Dahlberg A., Garrett R.A., Moore P.B., Schlessinger D., and Warner J.R. (eds). *The Ribosome: Structure, Function, and Evolution*. American Society for Microbiology, Washington D.C. pp148-159.
- Noller H.F., Moazed D., Stern S., Powers T., Allen P.N., Robertson J.M., Weiser B., and Triman K. (1990). In Hill W.E., Dahlberg A., Garrett R.A., Moore P.B., Schlessinger D., and Warner J.R. (eds). *The Ribosome: Structure, Function, and Evolution*. American Society for Microbiology, Washington D.C. pp73-92.
- De Rijk P., Neefs J-M., Van de Peer Y., De Wachter R., (1992). *Nucleic Acids Res.* 20, Supplement 2075-2089.
- Gutell R.R., Power A., Hertz G.J., Putz E.J., and Stormo G.D. (1992). *Nucleic Acids Res.* 20, 5785-5795.