The EMBO Journal vol.5 no.5 pp.1111-1113, 1986

Higher order structure in ribosomal RNA

R.R.Gutell, H.F.Noller[1] and C.R.Woese

Department of Genetics and Development, University of Illinois, Urbana, IL 61801, and [1]Thimann Laboratories, University of California at Santa Cruz, Santa Cruz, CA 95064, USA

Communicated by R.A.Garrett

The only reliable general method currently available for determining precise higher order structure in the large ribosomal RNAs is comparative sequence analysis. The method is here applied to reveal 'tertiary' structure in the 16S-like rRNAs, i.e. structure more complex than simple double-helical, secondary structure. From a list of computer-generated potential higher order interactions within 16S rRNA one such interaction considered likely was selected for experimental test. The putative interaction involves a Watson-Crick one to one correspondence between positions 570 and 866 in the molecule (*E. coli* numbering). Using existing oligonucleotide catalog information several organisms were selected whose 16S rRNA sequences might test the proposed co-variation. In all of the (phylogenetically independent) cases selected, full sequence evidence confirms the predicted one to one (Watson-Crick) correspondence. An interaction between positions 570 and 866 is, therefore, considered proven phylogenetically.

Key words: ribosomal RNA/comparative sequence analysis

Introduction

The secondary structural elements in the 16S-like ribosomal RNAs proposed on the basis of sequence comparisons (Woese *et al.*, 1980; Woese *et al.*, 1983; Gutell *et al.*, 1985) are now for the most part supported by an overwhelming amount of evidence. One or more base pair replacements have been noted for every (or most) position(s) in many of the proposed structures (Gutell *et al.*, 1985; unpublished results). The results of recent detailed chemical probing experiments are also consistent with the proposed helices (Moazed *et al.*, 1986). Higher order structure for the large rRNAs other than that in helices is another matter, however. Experience with tRNA suggests that relatively little variation in composition will be seen for most 'tertiary structural' elements (Kim, 1979) and given that such elements are generally single pairs (or triples), finding sufficient comparative evidence to demonstrate convincingly their existence should be difficult.

Results and Discussion

An alignment of the existing published 16S-like rRNA sequences (as of mid 1984) was screened by computer for co-varying positions (Gutell *et al.*, 1985). One of those for which fairly extensive evidence existed involved positions 570 and 866. The interaction in question is one of a few that can be traced in part in the large number of oligonucleotide catalogs (about 400) existing for 16S-like rRNAs. By this means other rRNAs whose sequences might serve to test the hypothetical interaction could be selected. Table IA shows the original information suggesting the co-variation between positions 570 and 866 (Gutell *et al.*, 1985).

For those cases in which it is a C (i.e. the eubacteria and most mitochondria), position 866 tends to be represented in catalogs by oligonucleotides of the form (G)YUAACR (Woese *et al.*, 1983; Y = pyrimidine, R = purine). A majority of the oligonucleotide catalogs in all the major eubacterial groups contain an oligonucleotide of the YUAACR type, except for the *Bacteroides* and relatives and the *Planctomyces* — in which cases such a sequence has never been found (Woese *et al.*, 1985). Similarly, cases in which position 570 is a pyrimidine tend to be represented by oligonucleotides of the form (G)YYUAAAG (Woese *et al.*, 1983, 1985). All archaebacterial sequences and catalogs contain an oligonucleotide of the form CYUAAAG. These are rarely found in eubacterial catalogs, however, with the exception again of the *Bacteroides* and relatives and the *Planctomyces* — wherein all catalogs do contain such a sequence. Furthermore, in the archaebacterial catalogs a mutually exclusive relationship exists between CCUAAAG and CUUAAAG (Woese *et al.*, 1984; unpublished data).

On this information 16S rRNAs of eubacteria from the bacteroides group (e.g. *Bacteroides fragilis*) or the planctomyces group (e.g. *Planctomyces stayleyi*) and certain archaebacteria (e.g. *Methanospirillum hungatei*, *Sulfolobus solfataricus* and *Thermoproteus tenax*) were selected to test the predicted co-variation. The composition of positions 570 and 866 in these sequences is shown in Table IB. In all cases tested (Weisburg

Table I. Sequence of various 16S rRNA in the vacinity of the 570/866 co-variation. Positions 570 and 886 are in capitals, the remaining positions in lower case; y = pyrimidine, r = purine, n = any base

A. Original sequences suggesting co-variation (Gutell *et al.*, 1985)

570	866	
:	:	
gc G uaaag	gcyaa C gcg	eubacteria (four), chloroplasts (three), plant mitochondria
gc G uauag	gcuaa C gca	*Mycoplasma capricolum*
gc G uaaag	guuaa C acg	chloroplast *Euglena*
gc G uacag	guuaa C aca	ciliate mitochondrion
gc C uaaag	nnnaa C aa.	mammalian mitochondria (four)
gu U uaaag	augaa A rug	fungal mitochondria (two)
gc C uaaag	gggaa G ccg	archaebacteria (three)
ag U uaaaa	grgaa A yca	eucaryotes (six)

B. Additional sequences testing co-variation

570	866	
:	:	
gu U uaaag	gcgaa A gca	*Bacteroides fragilis* (Weisberg *et al.*, 1985)
gc U uaaag	gcgaa A gug	*Planctomyces staleyi* (Oyaizu and Woese, in preparation)
gc U uaaag	gggaa A ccg	*Methanospirillum hungatei* (Yang *et al.*, 1985)
gc C uaaag	gggaa G cyg	*Sulfolobus solfataricus* (Olsen *et al.*, 1985)
gc U uaaag	gggaa A ccg	*Thermoproteus tenax* (Leinfelder *et al.*, 1985)

et al., 1985; Yang *et al.*, 1985; Leinfelder *et al.*, 1985; Olsen *et al.*, 1985; Oyaizu and Woese, in preparation) the predicted co-variation between positions 570 and 866 holds. Thus, within the eubacteria and mitochondria two versions of the correspondence (G/C and U/A) exist, and their phylogenetic distribution demands that one of them (probably U/A) has arisen at least twice, if not three times. The archaebacteria show the two versions C/G and U/A, and their phylogenetic distribution again requires at least two independent occurrences for one of the versions. Eucaryotes seem to exhibit only the U/A version of the correspondence. Many new 16S-like rRNA sequences have been deduced since 1984 (when the initial computer screen was carried out), and none of them violate the Watson–Crick correspondence. We feel the hypothetical relationship between position 570 and 866 in 16S-like rRNAs has now been reliably demonstrated.

Positions 570 and 866 would seem to be involved in rather complicated interactions. The structures involved and the comparative evidence for them are as follows. Position 866 is the final nucleotide in a loop of four bases defined by a double-stranded stalk of two canonical pairs. Evidence for the existence of these pairs (861–2/867–8) is convincing. They are GY/RC in the eubacteria and in the plant and ciliate mitochondria, AU/RU in fungal mitochondria, UU/AA in the mouse mitochondrion, GG/CC in archaebacteria, and GA/UC or GC/GC in eucaryotes (Gutell *et al.*, 1985; unpublished data). The eubacterial oligonucleotide catalogs show a number of phylogenetically independent occurrences of U/A pairs at 862/867 replacing C/G pairs (Woese *et al.*, 1983). No proven counter-examples to these canonical pairing co-variations exist.

The possibility for base pairing also exists between positions 571 and 865 (immediately adjacent to 570/866), but in all cases the pair has the invariant composition U/A (Gutell *et al.*, 1985; unpublished data).

Another structure that could further constrain geometry in this general area of the 16S rRNA is a helix 567–569/881–883. Comparative evidence exists for the 569/881 pair. In most eubacteria, all archaebacteria and most mitochondria its composition is C/G, in eucaryotes G/C, but in one eubacterial group (the bacteroides) and in the fungal mitochondrial examples it is U/A (Gutell *et al.*, 1985; Weisburg *et al.*, 1985). Some comparative evidence exists for the 567/883 pair as well. In eubacteria, most mitochondria and archaebacteria it is G/C; the ciliate mitochondrion has A/U; and eucaryotes C/G (Gutell *et al.*, 1985; unpublished data). There is insufficient comparative evidence to claim the intervening pair, 568/882, as proven phylogenetically.

The structures involved in and surrounding the potential higher order interactions are shown in Figure 1. If one assumes that all these helices exist simultaneously and are maximally stacked, contradiction occurs. It is sterically impossible for a pair between positions 570 and 866 to stack on top of both the 861–2/867–8 helix and the 567–9/881–3 helix simultaneously. The fact that each of these three possible helical structures seems to involve only 2–3 bp strongly suggests each has to be supported by surrounding structure. The possible steric interactions are too many and too complex to sort out on the basis of present information.

While the existence of an interaction between positions 570 and 866 seems certain and will serve to constrain significantly the overall geometry in one particular locale of the molecule, the stereochemistry of this higher order structure remains uncertain.

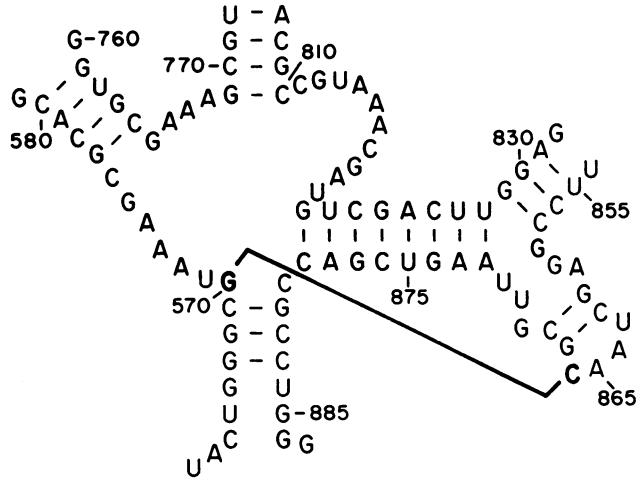

Fig. 1. Secondary structural elements bearing upon the 570/886 interaction. See text for details.

It is important to emphasize that any projected higher order structural interactions must be accompanied by substantial supporting evidence. If strict criteria are not used, the large number of incorrect interactions proposed under more relaxed conditions would serve to obscure or discredit the true ones suggested. This point, valid enough for the smaller tRNAs and 5S rRNA, is doubly so for the larger and much more complex ribosomal RNAs.

A number of higher order interactions have already been proposed that do not withstand rigorous comparative test. Thompson and Hearst (1983) have suggested three of them for the 16S rRNA, while Spitnik-Elson *et al.* (1985) have proposed 17. Comparing the first group of proposed structures in a set of 22 eubacterial sequences that has been properly aligned shows the vast majority of the nucleotide replacements to result in mispairs. For example, each of the seven positions in the putative 620–626/1520–1426 helix (Thompson and Hearst, 1983) shows two or more mismatches. The few properly compensated bp replacements that do result are within chance expectation. Similarly seven or more of the tertiary interactions proposed by Spitnik-Elson *et al.* (1985) show a relatively high level of mispairs when tested on a properly aligned and extensive set of eubacterial sequences (unpublished data).

Acknowledgements

This work is supported by grants from the National Aeronautics and Space Administration, NSG-7044 (to C.R.W.), and from the National Institutes of Health, GM-17129 (to H.F.N.).

References

Gutell,R.R., Weiser,B., Woese,C.R. and Noller,H.F. (1985) In Cohn,W.E. and Moldave,K. (eds), *Prog. Nucleic Acid Res. Mol. Biol.*, **32**, 155–216.

Kim,S.-H. (1979) In Schimmel,P., Söll,D. and Abelson,J. (eds), *Transfer RNA: Structure, Properties, and Recognition.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 83–100.

Leinfelder,W., Jarsch,M. and Böck,A. (1985) *Syst. Appl. Microbiol.*, **6**, 164–170.

Moazed,D., Stern,S. and Noller,H.F. (1985) *J. Mol. Biol.*, **187**, 399–416.

Olsen,G.J., Pace,N.R., Nuell,M., Kaine,B.P., Gupta,R. and Woese,C.R. (1985) *J. Mol. Evol.*, **22**, 301–307.

Paster,B.J., Ludwig,W., Weisburg,W.G., Stackebrandt,E., Hespell,R.B., Hahn,C.M., Reichenbach,H., Stetter,K.O. and Woese,C.R. (1985) *Syst. Appl. Microbiol.*, **6**, 34–42.

Thompson,J.F. and Hearst,J.E. (1983) *Cell*, **32**, 1355–1365.

Spitnik-Elson,P., Elson,D., Avital,S. and Abramowitz,H. (1985) *Nucleic Acids Res.*, **13**, 4719–4738.

Weisburg,W., Oyaizu,Y., Oyaizu,H. and Woese,C.R. (1985) *J. Bacteriol.*, **164**, 230–236.
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. and Noller,H.F. (1980) *Nucleic Acids Res.*, **8**, 2275–2293.
Woese,C.R., Gutell,R., Gupta,R. and Noller,H.F. (1983) *Microbiol. Rev.*, **47**, 621–669.
Woese,C.R., Gupta,R., Hahn,C.M., Zillig,W. and Tu,J. (1984) *Syst. Appl. Microbiol.*, **5**, 97–105.
Woese,C.R., Stackebrandt,E., Macke,T.J. and Fox,G.E. (1985) *Syst. Appl. Microbiol.*, **6**, 143–151.
Yang,D., Kaine,B.P. and Woese,C.R. (1985) *Syst. Appl. Microbiol.*, **6**, 251–256.

Received on 14 January 1986