Volume 11 Number 15 1983 Nucleic Acids Research

5S rRNA sequences from four marine invertebrates and implications for base pairing models of metazoan sequences

William F.Walker and W.Ford Doolittle

Department of Biochemistry, Dalhousie University, Halifax, Nova Scotia B3H 4H7, Canada

Received 25 April 1983; Revised and Accepted 12 July 1983

ABSTRACT

The nucleotide sequences of 5S rRNAs from the starfish Asterias vulgaris, the squid Illex illecebrosus, the sipunculid Phascolopsis gouldii and the jellyfish Aurelia aurita were determined. The sequence from Asterias lends support for one of two previous base pairing models for helix E in metazoan sequences. The Aurelia sequence differs by five nucleotides from that previously reported and does not violate the consensus secondary structure model for eukaryotic 5S rRNA.

INTRODUCTION

The sequences of 5S rRNAs from a wide variety of metazoans have been published during the past two years (1), partly as a probe of phylogenetic relationships within this group of organisms. Here, we report four additional sequences from marine invertebrates, three of which represent classes or phyla not previously reported on. A sequence was recently determined for one of these species (Aurelia aurita; ref 2). However, as there are significant differences in the two sequences from this species, we include our sequence. We discuss the implications of one of our sequences for base pairing models for helix E in metazoa.

MATERIALS AND METHODS

RNA was isolated from the ovaries of the starfish Asterias vulgaris and the jellyfish Aurelia aurita, from the digestive gland of the squid Illex illecebrosus and from the coelomic fluid of reproductively active sipunculids, Phascolopsis gouldii. The animal tissues or fluids were suspended in isolation medium (3) and macerated by a Polytron^{R} at full speed for 30 sec. before extraction with buffer-saturated phenol. 5S rRNA was purified (4) and the 3' or 5' termini were labelled with ^{32}P (5). Partial digests of 3'-labelled material, using the chemical sequencing method of Peattie (6) and the enzymatic method of Donis-Keller (7) were analyzed on polyacrylamide gels

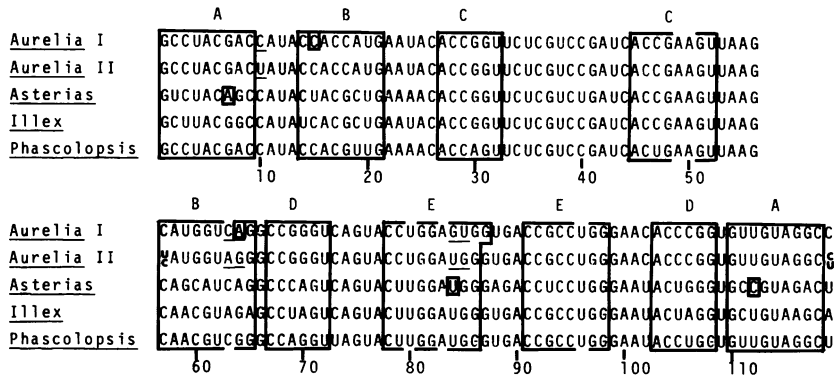

Fig. 1. The 5S rRNA sequences from the jellyfish Aurelia aurita as previously reported (I) and as herein determined (II), from the starfish Asterias vulgaris, the squid Illex illecebrosus and the sipunculid Phascolopsis gouldii. The proposed helical regions according to the model of Böhm et al. (8) are boxed, and those positions in which the two sequences from Aurelia aurita differ are underlined.

(3). Terminal nucleotide analyses were also performed (5).

RESULTS AND DISCUSSION

The sequences are aligned with each other and with the sequence from Aurelia aurita previously determined (Fig. 1; ref 2). The helical regions according to the consensus model for eukaryotic 5S rRNAs (8-10) are indicated and the five positions in which the two Aurelia aurita sequences differ from each other are underlined.

The five nucleotide differences between the two Aurelia aurita sequences could conceivably represent a real intraspecific divergence between widely separated source populations. However, we note that two unprecedented deviations from the consensus secondary structure model for eukaryotic 5S rRNAs are evident in the sequence from Hori et al., (2) but are not present in the sequence here reported (Figs. 1, 2 and 3 a-c). Firstly, an A-C mispair in the second base pair of helix B is implicit in the former sequence. To our knowledge, a Watson-Crick base pair is present in the first two base pair positions, beyond the looped out nucleotide, of helix B in all other eukaryotic 5S rRNAs. Secondly, three rather than the usual four nucleotides are present in the hairpin loop of helix E (Fig. 3c). The sequence of Hori et al. (2) differs from ours in this region only in an apparent reversal of the G and U in positions 84 and 85. There was no question of the ordering of these two

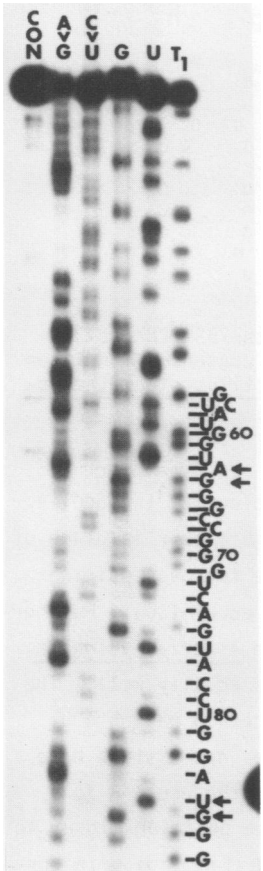

Fig. 2. X-ray photograph showing part of the 5S rRNA sequence from the jellyfish Aurelia aurita as determined by partial chemical digests (6). con = control, A = adenosine, C = cytosine, G = guanosine, U = uracil, T_1 = T_1 ribonuclease (guanosine-specific). Positions where our sequence differs from that of Hori et al. (2) are indicated by an arrow.

nucleotides on our gels (Fig. 2). Another difference between the two Aurelia aurita sequences involves a U vs. a C in position 10. C is present here in all other eukaryotic sequences except for one strain of Chlamydomonas reinhardtii, which also has a U in this position (1). In addition to the 5 full differences between the two Aurelia aurita sequences, we detected nucleotide heterogeneity in positions 57 and 119.

Currently, two base-pairing models for helix E (IV of some authors) in metazoa are favored by most authors. Model A (Fig. 3a), first proposed by Komiya et al. (11), is characterized by a single nucleotide loop out in a position corresponding to a similar loop out in all other eukaryotic 5S sequences (1). A U-U mispair is invariably present 3 pairs from one end of the helix. Model B (Fig. 3b) has gained wide acceptance recently (1,2,9,12). A C-A mis-

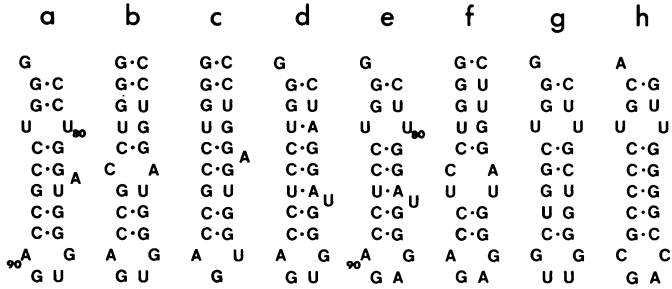

Fig. 3. Alternative base pairing models for helix E of the 5S rRNAs of Aurelia aurita (a-c) and Asterias vulgaris (e,f) and the consensus model for wheat (Triticum aestivum) (d), Prochloron sp. (g) and Bacillus stearothermophilus (h).

pair is invariably present in place of the looped out nucleotide in Model A. No U-U mispair is present; however, there is typically a G:U pair in place of a G:C. Thus, the stability of helix E according to these two base pairing schemes has generally been concluded to be approximately equal (10). Furthermore, all known metazoan sequences, except for those from Asterias vulgaris, (Fig. 3) and Lineus geniculatus (13) fit these two models equally well. The Asterias sequence differs from all other metazoan sequences in having a U rather than a G in position 93 (Fig. 1 and 3e). This does not disturb base pairing according to model A, except that a U rather than the usual A is looped out and the position of the loop out is shifted one position closer to the hairpin, such as occurs in sequences from angiosperms (Fig. 3d). In contrast, according to Model B, a U-U mispair occurs adjacent to the C-A mispair in this sequence (Fig. 3f). Thus, there are two fewer Watson-Crick base pairs in the helix, according to model B. Also, as previously indicated (13), the sequence from the ribbon worm Lineus geniculatus has G-G and C-U mispairs in addition to an A-C mispair in Helix E, according to model B. Here, we point out that this sequence fits model A well, with only a U-U mispair in the usual position. We further emphasize that U-U mispairs are apparent in certain sequences from prokaryotes (e.g. Fig. 3 g,h) in the same position indicated in model A. Finally, U-U mispairs in the penultimate base pair position of helix A are apparently universal and of ancient origin in the major basidiomycete lineage (ref. 14 and unpublished data). Thus, we feel that the total sequence data from metazoans, prokaryotes and basidiomycetes supports the proposal that U-U mispairs in certain positions within 5S rRNA helices can be evolutionarily stable alternatives to Watson-Crick or G:U base pairs.

Table 1. Matrix of nucleotide differences between selected 5S rRNA sequences from metazoa.

	I.i.	H.p.	M.e.	L.a.	A.v.	L.v.	A.a.	H.p.	P.g.	D.m.	H.s.
Illex illecebrosus	-	22	17	18	20	21	18	15	17	23	23
Helix pomatia	22	-	13	16	26	22	23	16	17	17	27
Mytilus edulis	17	13	-	7	19	18	21	16	10	19	21
Lingula anatima	18	16	7	-	18	21	19	17	9	22	20
Asterias vulgaris	20	26	19	18	-	21	24	20	21	26	21
Lytechinus variegatus	21	22	18	21	21	-	18	20	16	18	21
Aurelia aurita	18	23	21	19	24	18	-	14	14	20	23
Halichondria panicea	15	16	16	17	20	20	14	-	12	22	21
Phascolopsis gouldii	17	17	10	9	21	16	14	12	-	18	19
Drosophila melanogaster	23	17	19	22	26	18	20	22	18	-	25
Homo sapiens	23	27	21	20	21	21	23	21	19	25	-

A matrix of nucleotide differences between the sequences herein reported and selected metazoan sequences (Table 1) provides little encouragement that 5S sequences will make a significant contribution toward an understanding of many of the phylogenetic relationships between metazoan phyla and between the classes of molluscs and echinoderms. A comparison of the sequences from molluscs suggests a more recent divergence of the sequences from the gastropod Helix and bivalve Mytilus than from the sequence of the cephalopod Illex illecebrosus. However, the rather extensive molluscan fossil record is not compatable with this result (15). Also note that the sequence from the brachiopod Lingula anatina is clearly more similar to that from the bivalve Mytilus than are the sequences from Helix and Illex. Yet the fossil record of bivalves and brachiopods begins in the early Cambrian when the major radiation of present metazoan phyla is believed to have occurred (16). Thus, the unexpected degree of similarity of the sequences from Mytilus, Lingula and Phascolopsis appears to be more the result of relatively slow rates of substitutions and perhaps convergence rather than an indication of a more recent common ancestor. A sufficiently large sample size from each higher taxon may cancel some anomalous phylogenetic inferences. However, it is probable that larger molecules with essentially qualitative markers such as insertions and deletions will be essential in providing a confident phylogenetic tree for the

metazoa.

Since completion of Table 1, 5S rRNA sequences were reported from the annelids Perineresis brevicirris and Sabellastarte japonica, the echiurid Urechis unicinctus and the nemertian Lineus geniculatus (13). Morphological and embryological criteria suggest a phylogenetic relationship between sipunculids, echiurids and annelids (17). Comparisons of the 5S rRNA sequence from the sipunculid Phascolopsis gouldii (Fig. 1) with the above sequences yields 9, 14, 6 and 11 different nucleotides, respectively. Thus, the sequence data is consistent with the general belief (17) that the two phyla of unsegmented coelomate worms, echiurids and sipunculids, are most closely related to each other.

ACKNOWLEDGEMENTS

We thank Murray Schnare for performing end nucleotide analyses, the Dorothy Killam Trust Fund for supporting WFW as a postdoctoral fellow, and the Natural Sciences and Engineering Research Council and Medical Research Council of Canada for support.

REFERENCES

1. Erdmann, V.A., Huysmans, E., Vandenberghe, A. and DeWachter, R. (1983) Nucl. Acids Res. 11, r105-r133.
2. Hori, H., Ohama, T., Kumaziki, T. and Osawa, S. (1982) Nucl. Acids Res. 10, 7405-7408.
3. Walker, W.F. and Doolittle, W.F. (1982) Nucl. Acids. Res. 10, 5717-5721.
4. Donis-Keller, H., Maxam, A.M.. and Gilbert, W. (1977) Nucl. Acids Res. 4, 2527-2538.
5. MacKay, R.M., Gray, M.W. and Doolittle, W.F. (1980) Nucl. Acids Res. 8, 4911-4917.
6. Peattie, D.A. (1979) Proc. Natl. Acad. Sci. USA 76, 1760-1769.
7. Donis-Keller, H. (1980) Nucl. Acids Res. 8, 3133-3142.
8. Bohm, S., Fabian, H. and Welfle, H. (1981) Acta Biol. Med. Germ. 40, K19-K29.
9. DeWachter, R., Chen, M.W. and Vandenberghe, A. (1982) Biochimie 64, 311-329.
10. Delihas, N. and Andersen, J. (1982) Nucl. Acids. Res. 10, 7323-7344.
11. Komiya, H., Shimizu, N., Kawakami, M. and Takemura, S. (1980) J. Biochem. 88, 1449-1456.
12. Kuntzel, H., Piechulla, B. and Hahn, U. (1983) Nucl. Acids Res. 11, 893-900.
13. Kumazaki, T., Hori, H. and Osawa, S. (1983) Nucl. Acids Res. 11, 3347-3350.
14. Walker, W.F. and Doolittle, W.F. (1982) Nature 299, 723-724.
15. Yochelson, E.J. (1979) in The Origin of Major Invertebrate Groups, House, M.R., Ed., pp. 323-358, Academic Press, N.Y.
16. Brasier, M.D. (1979) in The Origin of Major Invertebrate Groups, House, M.R., Ed., pp. 103-159, Academic Press, N.Y.
17. Meglitsch, P.A. (1972) Invertebrate Zoology, 2nd ed., Oxford University Press, London, p. 437.